Spatio-temporal modelling of under-five mortality and associations with malaria-anaemia comorbidity and health interventions in sub-Saharan Africa

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Isidoros Papaioannou

von Griechenland

2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Dr. Jürg Utzinger (Fakultätsverantworlicher), PD Dr. Penelope Vounatsou (Dissertationsleiter), und Prof. Dr. Anna-Sofie Stensgaard (Korreferent)

Basel, den 18. Februar 2020

Prof. Dr. Martin Spiess

Dekan

To my beloved parents, Dimitris and Vasiliki

Summary

A remarkable reduction of the total number of under-five deaths was achieved between 1990 and 2018 in the African setting, as pre-school mortality fell to 5.3 million deaths compared to 12.5 million in 1990. The bulk share of this reduction is attributed to the Millennium Development Goals (MDGs) era, during which time the under-five mortality rate has been declining with an annual rate of 3.8% across Africa. Despite these important achievements, the sub-Sahara African region did not meet the fourth target of the MDGs and still has an unacceptably high under-five mortality rate. Crucially, limiting the under-five mortality rate to a maximum of 25 deaths per 1,000 live births by 2030 lies at the heart of the Sustainable Development Goals (SDGs) and a recent report from the United Nations has warned that based on current trends, the African continent will not meet the SDG target for under-five mortality. Hence, providing useful insights from the associations between under-five mortality, the leading causes of disease and preventative or curative health interventions could make available valuable information to decision makers in order the African countries to achieve the SDGs on pre-school mortality.

Malaria is a major contributor to under-five mortality in sub-Saharan Africa, accounting for 400,000 deaths, approximately 60% of which are in children below the age of five. At global scale, the disability-adjusted life-years for the malaria disease are 45 million. An important aspect of the disease is that infection by malaria parasites does not necessarily lead to mortality and it is rather conditions that follow infection or other comorbidities that produce severe forms of the disease with increased mortality risk. Apart from malaria, pneumonia

and diarrhea account for the most frequent causes of pre-school deaths. An interesting feature of all these three leading causes of under-fives in Africa, i.e. pneumonia, diarrhea and malaria, is that they share febrile response as their main clinical manifestation. Against the leading causes of under-five mortality, preventative or curative health interventions have been widely adopted in Africa, with their spatial coverage being on a significant rise, particularly due to the so-called scaling-up of health interventions during the last five years of the MDGs. For instance, ownership of Insecticide-Treated nets against malaria rose from 50 to 80 percent between 2010 and 2015, while their utilization averted 663 million clinical malaria cases over the MDGs era. Yet, the coverage of health interventions and the subsequent reduction in under-five deaths has happened in an unequal way across sub-Saharan Africa, raising concerns about health inequities at sub-national level.

The overall aim of the present PhD thesis is to develop, implement and interpret Bayesian geostatistical models with spatially varying coefficients in order to analyze approximately one million, cross-sectional mortality related-data in Africa and associate under-five mortality with malaria and health interventions. The point-by-point objectives of our work are as follows:

1. To develop a novel indicator for quantifying malaria-related mortality for children under the age of five in sub-Saharan Africa, namely the malaria-anemia comorbidity prevalence indicator (chapter 2);
2. To identify health inequities experienced by sub-national populations due to the geographical variation in the association between curative or preventive health interventions and under-five mortality in sub-Saharan Africa (chapter 3);
3. To assess the contribution of the leading causes of under-five mortality in sub-Saharan Africa on febrile response by associating the prevalence of malaria parasitaemia, diarrhoea and ARI with fever. (chapter 4);
4. To estimate the association between health interventions and under-five mortality on changes in mortality risk between two time points across Africa (chapter 5);

5. To compare Bayesian variable selection methods for spatially varying coefficient models, given that these approaches are at the forefront of analyzing geolocated mortality data in Africa (chapter 6).

In chapter 2, we estimated the association of malaria parasitaemia, anemia, and malaria-anemia comorbidity with all-cause under-five mortality and evaluated the potential of malaria-anemia comorbidity prevalence to quantify malaria-related deaths in sub-Saharan Africa. Additionally, we estimated within-country variation of the association between comorbidity and under-5 mortality, using spatially varying coefficient models. We presented our results at high spatial resolution, including model-based risk maps of malaria, anemia, and malaria-anemia comorbidity.

In chapter 3, we modeled the geographical variation in the association between health interventions and all-cause, under-five mortality in order to identify health inequities experienced by sub-national populations within a given country. To achieve that, we developed Bayesian geostatistical Weibull survival models with spatially varying coefficients for the effect of health interventions on mortality. Our approach allowed us to calculate the number of statistically important associations between interventions and mortality at regional level and hence to assess if health equity of interventions exists across the regions of a given country.

In chapter 4, we assessed the contribution of the leading causes of under-five mortality in sub-Saharan Africa on febrile response by associating the prevalence of malaria parasitaemia, diarrhoea and ARI with fever. Our flexible Bayesian spatial modelling approach allowed evaluating the geographical distribution of disease-exposure effect on fever in space (Administrative level 1). We also calculated the Potential Attributable Fraction (PAF) in order to quantify the contribution of childhood diseases on fever.

In chapter 5, we developed a novel methodology to statistically model the effect of health interventions on the changes in under-five mortality risk between two DHS survey time-

points for 21 countries in Africa. We used a Bayesian geostatistical Weibull survival modeling approach and implemented rigorous Bayesian variable selection procedures in order to identify the most suitable set of health interventions for subsequent model fit.
In chapter 6, we assessed the performance of stochastic search variable selection (SSVS) for the fixed effects of geostatistical models, we compared three different Bayesian variable selection (BVS) methods for conditionally autoregressive (CAR) structured spatially varying coefficient models and finally we assessed the sensitivity of SSVS for the fixed effects when is co-implemented with a BVS procedure. We conducted a simulation study and applied the methods to the Burundi DHS in order to assess the aforementioned selection procedures.
The present PhD thesis has contributed to the scientific fields of Epidemiology and Statistics by committing to the spatio-temporal modelling of under-five mortality data in the African setting, using primarily routinely collected, cross-sectional, household-based survey data coming from the Demographic and Health surveys program. The key outcomes of the research conducted in this thesis are as follows:

1. Our work contributed to the development, proposal and validation of a novel indicator for quantifying malaria-mortality using survey data, i.e. the malaria-anemia comorbidity indicator. Our main conclusions were that malaria burden in sub-Saharan Africa is considerably underestimated when anemia in not taken into account and that the malaria-anemia comorbidity prevalence provides a useful measure of the malaria-related deaths;
2. We presented the first study to assess sub-national health inequities, across most countries in Africa, by employing a spatial statistical modelling approach and routinely collected survey data coming from the DHS and MIS. Our results demonstrated strong sub-national health inequities across various regions for 28 African countries;
3. Our estimates confirmed the strong contribution of diarrhoea and acute respiratory infection on febrile response and accounted only one out of five cases to malaria;

4. Our work concluded that the health interventions that are mostly associated with changes in all-cause, under-five mortality risk in sub-Saharan Africa were Bacillus Calmette–Guérin (BCG) immunization, vitamin A supplementation and deworming medication;
5. Our analysis showed that the SSVS method is able to accurately identify the statistically important predictors for the fixed effects of geostatistical models and that SSVS is not sensitive to co-implementation with a BVS procedure for CAR-structured, spatially varying coefficients. We also concluded that one of the three BVS methods for varying coefficients, namely the Global selection method, is able to identify true varying coefficients with 70% success rate.

Acknowledgements

As I write these words, I look back and reflect on all the people who helped and inspired me in many ways. While it is impossible to acknowledge everyone and every situation that led to this PhD thesis, I will try mentioning, to the best of my abilities, all those that offered a lot.

My deepest gratitude goes to my main supervisor, PD Dr. Penelope Vounatsou. Above all, I recognize and thank her for the unique opportunity that she provided me by offering a PhD position in her group. Without a doubt, Penelope is the expert on the intersection of statistics, epidemiology and spatial analysis. That said, Penelope generously shared her knowledge with me and showed me how research is conducted. I am most thankful for everything that she did for me.

My genuine thanks go to my co-supervisor and director of the Swiss TPH, Prof. Dr. Jürg Utzinger. Jürg is an inspiring person, a true leader and I would like to thank him for all the various ways he supported my research and all the motivational discussions we had.

My sincere thanks to Prof. Dr. Anna-Sofie Stensgaard for accepting the role of external examiner and assessing this PhD thesis.

I would like to express my warmest gratitude to all the present and past members of the Bayesian Biostatistics group for everything that we shared over the last four years. In no particular order, I would like to thank: Christos Kokaliaris, Anton Beloconi, Martin Matuska, Oliver Bäranbold, Carla Grolimund, Elizaveta Semenova, Julius Ssempiira, Sammy Khagayi, Ourohire Milogo, Betty Nambuusi and Salomon Massoda Tonye. Special thanks to Alex

Karagiannis-Voules for a much-needed mentorship at the beginning of my studies and all his support.

To all Swiss TPH staff: thank you for making this institute a stimulating environment for conducting research and for your support through the years. Special thanks to Natalie Wiedemar, Wendelin Moser and to every single person that made Swiss TPH a special place. My warmest wishes to all the lads from Swiss TPH football.

I am grateful to the European Research Council and the Swiss National Science Foundation for funding my thesis.

A BIG thank you to a few people that inspired me: Dr. Alexander Donev, Prof. Aristomenis G. Siskakis and Prof. George Tsaklidis.

I would like to express my warmest gratitude to a number of friends for their support through different stages of my life abroad: Ioanna Alifieraki, Panagiotis Kasimatis, Eleni Sotiropoulou, the Evangelinos brothers, Adrien Puledda, Elvira Carrio Gaspar, Nassim Rousset and Athanasios Kouklas. To my friends back home: Makis, Andreas, Giannis, Konstantinos, Achilles, Vasilis, Antonis, Kiki, Matina, Elli, Stavros and Yannis.

To my partner, Katerina: without you, it would have been impossible. Thank you for your endless love and support.

As a final note, I would like to thank my family: Dimitris, Vasiliki, Georgos, Eleni, Georgos, Despoina and Isidoros. These people gave their absolute everything to me and a few words in an acknowledgments section cannot depict their unprecedented help and support. It is to them that I dedicate this thesis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Under-five mortality

As an integral part of the Convention on the Rights of the Child, the United Nations acknowledged the inherent right that every child has to life and all its member states made a commitment to initiate all efforts needed for the maximum development and survival of children (UNICEF, 2019). To this regard, the under the age of five mortality rate (U5MR) is a crucial measure illustrating the health status of pre-school children within a society. The U5MR quantifies the probability for a child to die before its fifth birthday and it is presented as deaths per 1000 live births. The definition of live birth describes a conception product, which demonstrates evidence of being alive after having being expulsed or extracted from the mother (Woods, 2014). As of 2018, 5.3 million children worldwide did not survive their fifth birthday and the estimated global U5MR was 39 deaths per 1000 live births (WHO, 2019c).Yet pre-school mortality rates are unequally distributed across various regions. Specifically, sub-Saharan Africa (SSA) is the most afflicted region worldwide, accounting for the highest U5MR, i.e. 78 deaths per 1,000 live births and currently experiencing under-five mortality levels of developed countries approximately two decades ago (UN IGME, 2019). Crucially, the sub-Saharan region is home to the majority of countries exceeding an U5MR of 100 deaths per 1000 live births, with its children experiencing an estimated 8-fold chance of dying between birth and the age of five, compared to children of similar age residing in Europe.

Despite the unacceptably high, total number of preschool deaths and U5MR, a remarkable reduction of both measures was achieved between 1990 and 2018. Specifically, in 1990 the U5MR and the number of deaths were estimated at 93 per thousand live births and 12.5 million respectively, compared to an U5MR equal to 39 deaths and 5.3 million children that did not survive their fifth birthday in 2018. Importantly, the major decline of the above metrics was achieved between 2000 and 2015, hence during the era of the Millennium Development Goals (MDGs). The MDGs were eight goals established by the United Nations that aimed to improve global health and well-being, social justice, equality and environmental

sustainability. At the heart of these goals was the vision that in 2015, the U5MR would have been reduced by two-thirds, compared to 1990 (UN, 2019a). Despite the under-five mortality goal not being achieved, the millennium commitment contributed to a more than 50% decline in the U5MR. In sub-Saharan Africa, the U5MR fell from 182 deaths per thousand live births in 1990 to 78 in 2018. This translates into an annual reduction rate of 3% over these 28 years, albeit the highest annual rate was observed mostly during the MDGs era with an annual reduction rate equal to 3.8%. The global interest for a better, more sustainable life for all children was renewed with the introduction of the Sustainable Development Goals (SDGs), which now call for a U5MR of maximum 25 deaths per 1000 live births across all countries worldwide (UN, 2019b). Despite the considerable decline in U5MR in Africa over the past 28 years, a 2017 report from the United Nations projected future reductions in the U5MR and concluded that based on historical and current trends, the African continent will not meet the SDG target for under-five mortality (United Nations, 2017a). Additionally, under-five mortality has been unevenly reduced across various regions of Africa, with very few studies having assessed the subnational variation in mortality and its main drivers. Therefore, the leading causes of under-five mortality need to be studied, together with their spatial variation, in order to inform decision makers about the dynamics of mortality and diseases and hence to assist the African countries on their efforts to achieve the SDGs, in a fair way for all their populations. In addition to the major childhood diseases, other factors that influence under-five mortality, for which there is a priority to be studied and assessed at local-scale, are health interventions and socio-economic, demographic and environmental factors.

1.2 Childhood diseases

On a global scale, the distribution of childhood diseases that drive mortality for children less than five years of age is investigated separately for children in the first month of life (neonatal) and for those aged between one and 59 months. The reason behind this separation is that neonatal mortality accounts for a large share of all under-five deaths, i.e.

approximately 47% of all under-five mortality and since there are discrepancies between the leading causes of neonatal mortality and those for children aged 1 to 59 months (UN IGME, 2019). The major contributors to neonatal mortality are preterm birth complications, intrapartum-related events and sepsis. Conversely, pneumonia, diarrhoea, various injuries and malaria lead the causes of mortality for children between 1 and 59 months of age. In the African setting, neonatal deaths account only for 36% of total under-five mortality, albeit SSA has a neonatal mortality rate of 28 deaths per 1,000 live births, which is currently the highest globally. Additionally, due to the indigenous climatic and socio-economic conditions, tropical or poverty-related disease such as malaria and diarrhoea are more relevant in Africa than other parts of the world (WHO, 2018c). A supplementary but crucial factor underpinning under-five mortality is malnutrition, as it generally leads to weak immune systems that pave the way for easier development of diseases, and could additionally aggravate the type of disease that a child is already afflicted by. In Africa, it is estimated that malnutrition is associated with one-third of the total number of childhood deaths (WHO. Regional office for Africa, 2019f). Given its historical importance, significant burden and life loss, as well as its impact on socioeconomic development, malaria has a prominent role among the leading childhood diseases.

1.2.1 Malaria

1.2.1.1 Malaria characteristics and transmission

Parasitic, single-cell eukaryotes of the genus *Plasmodium* are responsible for the malaria disease. Human malaria is an outcome of a *Plasmodium* infection and specifically by the injection of *Plasmodium* sporozoites to an uninfected person. There are six *Plasmodium* species affecting humans, which are different in geographical distribution and in the severity of the malarial disease that they cause in humans (Cowman, 2016). The most prevalent *Plasmodium* species are *Plasmodium Falciparum* (*P. falciparum*) and *Plasmodium Vivax* (*P. vivax*), with *P. falciparum* claiming the most deaths. Although both species are present across different areas of the tropics and subtropics, *P. falciparum* is predominantly found in

sub-Saharan Africa while *P. vivax* in the Americas (Ashley, 2018). *Plasmodium malariae*, *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* are less prevalent than *P. falciparum* and *P. vivax* and are generally associated with milder forms of disease. *Plasmodium knowlesi* (*P. knowlesi*) is the last and most recently recognized of the six plasmodial species causing human malaria.

Female mosquitos of the genus *Anopheles* are the vectors responsible for the carriage and transmission of the *Plasmodium* parasites to humans. In particular, the *Anopheles* genus consists of approximately 460 species, over 100 of which have the capacity to transmit pathogens that can lead to human malaria (Wiebe, 2017). Yet, only about 40 species can successfully transmit the disease, as a mosquito species needs additional qualities, such as long life expectancy, presence in large numbers and parasite carrying capabilities, in order to sustain transmission (The Malaria Atlas Project, 2019). In the African setting, the *Anopheles gambiae* complex, consisting of several closely related, indistinguishable species and the *Anopheles funestus* species constitute the cornerstone of malaria transmission vectors. From the *Anopheles gambiae* complex, the *Anopheles gambiae sensu stricto* is considered among the most successful vectors of transmitting malaria, mainly due to its preference on feeding on humans (anthropophilic), the very short period in which its larva is developed and its generally long life span.

The malaria parasitaemia life cycle begins in the salivary grands of an infected mosquito, where cells, named sporozoites, capable of being infective agents when injected to humans, are developed. A blood meal of an infected mosquito will introduce the sporozoites to the dermis of the host. The sporozoites will move via the bloodstream into the liver, where they target and penetrate hepatocytes and within which they asexually multiply. This process results in the production of schizonts, which accommodate large numbers of hepatic merozoites. Rupture of schizonts allows the exit of merozoites from the liver and their later insertion into the circulatory system, in which the merozoites will use a four-stage process to penetrate red blood cells (RBC). Inside the RBC, the parasites de-differentiate into immature trophozoites and after a period of enlargement, they mature into schizonts. As before, the

schizonts are ruptured, resulting in the lysis of the RBC and the subsequent release of merozoites, which the schizonts contained. The released merozoites further penetrate uninfected RBC. The products of the rupture of RBC are the stimulants of the clinical manifestations of malaria. The continuation of malaria transmission occurs as some of the merozoites remain inactive and transform into micro- and macrogametocytes, which are later picked up by a mosquito during a blood meal. Inside the mosquito's gut, the gametocytes are transformed into gametes, which after fertilization form a zygote. Subsequently, the zygote forms an ookinete, which invade the gut walls and become an oocyst. The life cycle is concluded as sporozoites are produced within the oocyst and, after the rupture of the oocysts, the sporozoites end up in the mosquito's salivary glands (Tuteja, 2007; Soulard, 2015; Cowman, 2016; CDC, 2018a).

1.2.1.2 Clinical manifestations and malaria diagnosis

Acute febrile illness is typically the initial clinical sign of malaria infection, accompanied by other symptoms, resembling an Influenza-like illness (ILI). These symptoms differ among infected individuals, albeit usually including headache, joint pain, nausea, fatigue, chills and vomiting. The incubation period, representing the time interval between infection and initial clinical manifestation, spans from a mean of 12 days for *P. Knowlesi* and *P. falciparum*, to 17 days for *P. ovale* or *P. vivax* and to at least 18 days for *P. malariae* (Ashley, 2018). At this stage, malaria is characterized as uncomplicated and hence a timely, complete treatment can eliminate the infection. Any new, reoccurring sign of the symptoms could be due to unfinished treatment (*P. falciparum*) or resurgence of hypnozoites in the liver, which are products of *P. vivax* or *P. ovale* infection and inactively reside in the liver of the infected individual (Cowman, 2016).

Uncontrolled malaria infection resulted from only partial or complete absence of treatment for uncomplicated malaria, leads to severe malaria, which is a serious and potential fatal form of disease. The majority of severe malaria cases are due to *P. falciparum* infection and are dominated by damage to the central nervous system, for instance coma (cerebral malaria),

development of severe anemia and failures of the pulmonary or renal systems (Bartoloni, 2012). The aforementioned complications, with the addition of sudden bursting of an enlarged spleen, have been reported for the less frequent case of severe *vivax* malaria. Crucially, risk of severe malaria primarily concerns young children, as people living in endemic populations with stable transmission naturally acquire infection-immunity (premunition) to malaria, which means that an individual is protected against severe illness and high parasitaemia while still being blood infected.

Diagnosis of malaria concerns the observation of any clinical manifestations, the confirmation of the existence of parasites and the examination of an individual's health history. The gold standard for laboratory diagnosis for presence of parasites is the malaria (light) microscopy test. Specifically, the test utilizes blood films for identifying the presence of parasites, their density and their speciation. An alternative to microscopy testing is the Rapid Diagnostics Tests (RDTs) which use a patient's blood specimen extracted from a finger-prick to identify presence of malaria antigens and thus confirm parasitaemia (WHO, 2015a; CDC, 2018b). RDTs are currently extensively used in malaria-endemic countries, reaching an estimated high of 75% share among all malaria tests performed in SSA, against a 40% use in the year 2010 (WHO, 2019a). The rise of RDTs is mainly due to their simplicity in use, easy access to areas lacking good quality microscopy capabilities, speedy provision of an outcome and rather small amount of training required for the person performing the test.

1.2.1.3 Malaria determinants

Environmental factors directly influence malaria transmission and thus constitute an important malaria determinant. Temperature influences the time *Plasmodium* needs to develop inside the mosquito and below a specific threshold, it can even terminate the parasite cycle. Specifically, 18^0 C and 15^0 C are the absolute lowest temperatures at which *P. falciparum* and *P. vivax* can sustain their life cycle, while between 25^0 C and 30^0 C, the parasites experience their fastest development within the mosquito (Patz, 2006). Additionally, temperature influences the population dynamics of the *Anopheles* mosquito

itself, as for instance low temperatures extend the developmental period of mosquitos from eggs to adults and thus result in low mosquito densities (Beck-Johnson, 2013). Increase of altitude by 1,000 meters leads to a mean 6^0 C decrease in temperature and hence altitude influences malaria transmission and mosquito abundance (Patz, 2006). Rainfall also decisively influences the density of the *Anopheles* mosquitos, as it provides their breeding sites. As dry conditions influence the survival of the mosquitos, presence of scrubland or other suitable forms of vegetation may provide shelter and sustain transmission (Ricotta, 2014).

Socio-economic and demographic factors are major malaria determinants. Poverty and rural residency, for instance, are associated with lower housing quality, which allows the infected mosquitos to easier enter a residence and spread the disease (Tusting, 2015). Conversely, malaria disease may block individuals from regular business activities or substantially increase the expenses of families for diagnosis and treatment, thus leading to poverty. Knowledge of malaria characteristics, prevention capabilities and treatment solutions may be hindered by illiteracy or limited education of individuals (Degarege, 2019). Improved water and sanitation facilities are also associated with reduced malaria transmission (Bartram, 2005). Children under the age of five and pregnant woman constitute the most vulnerable groups becoming malaria infected.

1.2.1.4 Malaria burden

The global estimate for the number of malaria cases for the year 2017 is 219 million cases, reduced by 20 million cases from 2010 (WHO, 2018b). The Africa continent disproportionally shares 92% of all cases occurring worldwide, from which almost all (99.7%) are due to *P. falciparum*. Nigeria, Democratic Republic of the Congo and Mozambique account for the highest share of cases, followed by India and 12 more African countries. Globally, 435,000 deaths are attributed to malaria, approximately 60% of which are in children below the age of five. Similarly to the total number of cases, the African regions account for more than 92% of the total number of deaths worldwide. The disability-adjusted life-years (DALYs) for the

malaria disease at global scale are 45 million, encouragingly an approximately 34% reduction since 2007 (GBD 2017 DALYs and HALE Collaborators, 2018). An important aspect of malaria disease is fever, as it is the most usual clinical manifestation of parasitaemia infection and therefore malaria is directly associated with the fever burden.

1.2.2 Febrile response

Since childhood diseases are the causal agents of under-five mortality, an important safeguard against premature death is the observation and understanding of the clinical manifestations that a child exhibits. Febrile response constitutes one of the most frequent manifestations of illness and thus cause for medical consultation. It is defined as the above normal raising of the human body temperature as a response to an infection or inflammatory disease and which subsequently leads to improved protection through the stimulation of the immune system (Evans, 2015). The importance of fever as a clinical outcome is particularly relevant for the children below the age of five as this particular age band, together with elderly people and pregnant woman, have the highest risk of contracting diseases among the general population (WHO, 2019e). In the African setting, fever is a major public health concern as this part of the world accounts for the highest number of deaths under the age of five and because fever is the clinical manifestation of the leading causes of pre-school mortality, i.e. acute respiratory infection (ARI) presented as pneumonia, diarrhoea and malaria (WHO. Regional office for Africa, 2019f).

An important aspect of the manifestation of fever is to accurately identify which is the aetiology of a febrile response. In low and middle-income countries fever can be the outcome of a large set of potential diseases, for instance blood parasite diseases such as Malaria and Trypanosomiasis, viral infections such as Dengue fever or Influenza and bacterial zoonoses such as Q or Trench fever (Prasad, 2015). Additionally, the majority of low-income countries often lack the tools and health systems needed to identify the root cause of a febrile response in a patient, particularly in remote areas of a given country where access to diagnostic tools and health facilities is limited. Hence, febrile response is

sometimes wrongly attributed to a particular disease and subsequently falsely treated. In the African setting, children under the age of five presenting fever are often over-diagnosed and over-treated for malaria while, for instance, recent evidence has shown that only 37.7% of all malaria-positive under-fives experiencing fever had developed a febrile response due to *Plasmodium falciparum* infection (Dalrymple, 2019). These estimates are alarming, especially since recent studies have also suggested that health-seeking behaviour for fever among under-fives might be rather low, since less than half of pre-school children with febrile response would seek treatment, even if their home were less than a two-hours distance form a public health facility (Alegana, 2018).

1.3 Socio-economic, demographic and environmental factors

An analogous association of socioeconomic, demographic and environmental factors with malaria exists between these factors and under-five mortality (Yaya, 2018). The wealth index of the household in which a child resides has, for instance, a strong effect on U5MR, as children living in a poor environment may have only partial or even no access to crucial health interventions, such as malaria bednets, ACTs or antibiotics. Rural residence may also hinder child well-being due to limited access to health care facilities or logistical issues concerning the transportation and implementation of prophylactic or curative measures in remote areas. Maternal education is connected with greater knowledge of health-related issues and thus improved family responses to health challenges. Altitude, temperature and humidity are among the climatic factors associated with some of the leading causes of under-five mortality, such as respiratory infections, diarrhoea and malaria, and so contribute to mortality rate (Mirsaeidi, 2016; Liu, 2016; Azage, 2017).

1.4 Health interventions

Preventative or curative health interventions constitute the core tool against child mortality and consecutively their distribution and utilization play a major role for African countries in their efforts to achieve the 3.2 target of the SDGs. Such interventions can be broadly categorized into different groups, based on the means and against which diseases they aim

to provide prophylactic capabilities. Vaccination interventions are preventive health solutions that offer immunization against infectious diseases, among others, neonatal tetanus, poliomyelitis and measles. Malaria preventive interventions aim to minimize the vector-human interaction and include the access and utilization of Insecticide-treated nets (ITNs) and Indoor Residual Spraying (IRS) of households. Another example is micronutrient supplementation against Vitamin A or iron deficiency, which ensures child growth and provides support against infections (WHO, 2019d). In terms of curative solutions, prescription of antibiotics for children with ARI symptoms, Artemisinin-based combination therapy (ACT) against malaria, oral rehydration solution against diarrhoea and deworming medication are some of the core tools for treating the leading causes of under-five mortality in SSA.

The core interventions against malaria infection and malaria-related mortality can be classified into four different groups: vector control interventions, chemoprevention, as well as diagnosis and treatment (WHO, 2015b). Indoor Residual Spraying (IRS), Insecticide-Treated Nets (ITNs) and Long-Lasting Insecticide-Treated Nets (LLINs) are members of the vector control interventions group aiming to minimize the human-vector interaction. These interventions have played a crucial role in the prevention of an estimated 663 million cases of clinical malaria since 2000, 68% of which have been contributed to ITNs (Bhatt, 2015). Chemoprevention mainly refers to the use of intermittent preventive treatment (IPT) for pregnant women and infants in at least moderate malaria transmission areas of SSA and currently primarily focuses on *falciparum* malaria. Artemisinin-based combination therapy (ACT) is the current gold standard of *falciparum* and *vivax* malaria treatment, albeit for *vivax* malaria, treatment with chloroquine in non-chloroquine-resistant areas is possible (WHO, 2018a). In addition to the prior core interventions and as of now, the first malaria vaccine RTS,S/AS01 (RTS,S) is currently under evaluation for potentially being added to the core tools against malaria (WHO, 2019b). The vaccine is currently undergoing a pilot implementation programme in Ghana, Malawi and Kenya. As a final note, an ongoing threat to the core interventions is the emergence of insecticide resistant mosquitos, which results in

reduced efficacy of vector control interventions (Hemingway, 2016; Hancock, 2018; Kleinschmidt, 2018).

Following the Millennium commitment for improved child health and well-being, coupled with a renewed interest to fight leading causes of mortality such as malaria, most countries in Africa experienced a rapid scaling-up in the coverage of preventative or curative health interventions (Snow, 2015). As an example, ownership of ITNs increased from 50 to 80 percent over the last five years of the MDGs period, while access to ITNs within a household doubled (WHO, 2017a). The impact of these interventions on the survival of under-fives has been immense. For instance, it has been estimated that malaria control tools, championed by the use of ITNs, have averted 663 million clinical malaria cases for the period 2000-2015 (Bhatt, 2015). It has also been well-documented that vaccinations, improved hygiene facilities and oral rehydration solutions have assisted in the decline of pneumonia and diarrhoea-related deaths (UN IGME, 2019). Yet, despite the proven impact of such interventions, an important issue regarding their distribution is the significant discrepancies in the coverage of interventions within and between countries (Yourkavitch, 2018). This can be attributed to the existence of sub-national hotspots in many countries where coverage is notably lower compared to other areas. As differences in health intervention coverage may have a significant impact on the survival of children, it is of major public health importance to examine and prioritize potential sub-national health inequalities. The aforementioned issue is especially relevant for the SSA region, as this part of the world experiences the highest under-five mortality burden, coupled with generally low socio-economic status that further fuels inequality (UNICEF, 2014).

1.4.1 Health equities

The remarkable reductions in the total number of under-five deaths and the strong impact of health interventions on children's survival are two quantities that have not been equally distributed across space on the African continent. As of 2016, 13 SSA countries had an U5MR below 50 deaths per 1000 live births, while simultaneously seven countries had a rate

of more than 100 deaths (GBD 2016 Mortality Collaborators, 2017). Other studies have estimated and mapped at high spatial resolution unequal declines of under-five mortality across various regions in Africa over the MDGs period (Golding, 2017). For instance, large areas in Mali, Burkina Faso, Sierra Leone and the north of Nigeria and Cameroon still have an unacceptably high U5MR, whereas Rwanda, Uganda, Tanzania, Zimbabwe and Botswana saw important declines in mortality between 2000 and 2015. These discrepancies in U5MR can be partially explained by the variation in the effectiveness of health interventions. For instance, modelling studies utilizing cross-sectional, national and subnational representative, household-based survey data from Burkina Faso and Uganda concluded on strong sub-national variation in the associations between under-five mortality and child, maternal and household health interventions (Millogo, 2019; Nambuusi, 2019). It follows that these spatial discrepancies in the association of health interventions with U5M create inequalities for populations residing in areas where health solutions are less effective, compared to other regions within a country or a national average effect. The definition of health inequities describes inequalities that lead to unfairness and since these spatial variations in the effectiveness of health interventions incite unfairness for the people experiencing them, inequities in health interventions constitute an imminent chief priority for public health. Currently, modelling studies associating health interventions and under-five mortality at sub-national level are rather scarce, while at the same there is a sparsity of studies identifying health inequities based on the effectiveness of health interventions on pre-school mortality. Therefore, there is a need for modelling studies to assess the variation of the effect of health interventions in space, in order to provide key information to decision makers that will help to improve health equity in SSA.

1.5 Spatial variation of under-five mortality

The child mortality outcome is correlated in space, because it is influenced by environmental, socio-economic and other factors that are spatially varying. Geostatistical

models provide a state-of-the-art, flexible framework, which can incorporate such correlations.

1.5.1 Geostatistical modelling

Statistical models constitute a subclass of mathematical models, which quantify associations between random variables via mathematical expressions. Typically, one of the random variables called dependent variable is assumed to be explained by all others, namely independent variables, and it is assigned a probability distribution that is hypothesized to generate its realizations. Subsequently, one of the parameters of the distribution, assigned to the dependent random variable, is re-parametrized in terms of a function that contains a linear equation of all independent variables, i.e. the linear predictor. Assuming that the linear predictor contains only non-random quantities, i.e. the independent random variables for which we have observed data values, the model is called fixed effects model. An addition of random terms in the linear predictor, for which there are no observed data but are rather generated by a probability distribution, render the model a mixed effects model.

One of the most common assumptions of statistical modelling is the independence of observations. Yet this hypothesis is not appropriate for ecological studies and data derived from processes that inherit spatial and temporal correlations (Cressie, 2009; Hoeting, 2009). Instead, geostatistical models refer to a class of mixed-effect models that are employed under the assumption of a spatially correlated outcome. The term geostatistical refers to the type of geographical information we possess on the data. Specifically, the geographical information reflects fixed spatial points, for instance, coordinates, over a continuous study region, e.g. a country. Other kinds of spatial data than the geostatistical (point-level) include aggregated quantities over a spatial area (areal data) and observations for which their locations are derived from a random process (point pattern data). The spatial correlation in geostatistical models is usually incorporated in the linear predictor as a random term, which is assumed to follow a multivariate distribution that incorporates a functional form of some measure of the distance between unique data locations in its covariance matrix (Banerjee,

2014). Typically, a zero-mean multivariate Gaussian distribution with exponential correlation function is assumed. The introduction of geostatistical modelling was made by Diggle *et al.* in 1998 (Diggle, 1998) and since then these procedures have been extensively utilized for modelling associations, together with mapping efforts, of various diseases and their risk factors in the African setting. For instance, recent studies include the modelling of the malaria-mortality association in Kenya (Khagayi, 2019), the impact of changes in climate on malaria incident cases between 2013 and 2017 in Uganda (Ssempiira, 2018) and on spatial patterns of schistosomiasis in Zimbabwe (Pedersen, 2017).

The powerhouse of geostatistical modelling is Bayesian inference, as it allows for a more flexible, computationally advantageous model formulation over the frequentist approaches and addresses better heavy parametrized models for which frequentist approaches, such as the maximum likelihood estimation, are impractical (Gelfand, 1990). Bayesian inference comes in different flavours but its cornerstone has been the well-established, sampling-based Markov Chain Monte Carlo (MCMC) methods (McElreath, 2015). The MCMC procedures are based on repeated sampling from the posterior distribution of the model parameters, with the Gibbs sampling and its by-products being perhaps the most widely used MCMC techniques (Gelfand, 1990). More recently, efforts on Hamiltonian Monte Carlo have provided a faster and more effective procedure for MCMC (Gelman, 2013). Yet geostatistical modelling using MCMC can be computationally expensive or even infeasible, as the calculations involving the geostatistical intercept require matrix operations of dimensionality equal to the number locations in which the data were observed. A recent, reliable alternative to MCMC that particularly focuses on spatio-temporal data is a deterministic algorithm named Integrated Nested Laplace Approximation (INLA). The INLA procedure is based on Gaussian Markov random fields and the use of stochastic partial differential equations in order to accelerate all matrix operations concerning the geostatistical intercept (Rue, 2009; Lindgren, 2011).

1.5.2 Geostatistical variable selection

Given the correlations among the child mortality determinants and the spatial correlation in the under-five mortality outcome, geostatistical variable selection methods can assist on assessing the most suitable predictors for final model fitting between under-five mortality and a set of predictors. Variable selection methods are statistical procedures that evaluate the suitability of dependent variables for subsequent model-fit and prediction, based on some well-defined criteria that each method sets. When properly implemented, variable selection procedures can offer faster statistical analysis and improve overall prediction accuracy (Reunanen, 2003). For fixed effect models, Stochastic Search Variable Selection (SSVS) was the pioneering method for parameter selection in the Bayesian inference context, by enabling an inclusion indicator in the prior distribution (George, 1993). Following the work of George and McCulloch (1993), a series of different selection methods were proposed, for instance methods adapting the SSVS (Kuo, 1998; Dellaportas, 2002) or methods exploring simultaneously the whole of model space (Green, 1995), with O'Hara and Sillanpää making a recent review on such methods in the Bayesian context (O'Hara, 2009). Bayesian variable selection methods have been further extended to mixed effects model, for instance in logistic regression (Kinney, 2007; Yang, 2011). Recently, interest has been attracted to variable selection methods for spatially varying coefficient models, a specific branch of mixed-effects spatial models that, in addition to a geostatistical intercept explaining the correlation in the response, allow one or more coefficients to vary in space (Reich, 2010; Boehm Vock, 2015; Zhang, 2016; Choi, 2016).

1.5.3 Data sources

National civil registration and vital statistics systems (CRVS) constitute the data powerhouse that provides key information for the health status of a country and thus is the source for evidence-based policies concerning society and health. CRVS records events such as birth and death of an individual, causes of illness and mortality, as well as marriage-related measures (Ye, 2012). While prominent in developed nations, SSA countries completely lack,

or have only partially-developed CRVS systems, which renders the adoption of evidence-based policies challenging. To tackle this issue, Health and Demographic surveillance systems (HDSS) were established in SSA, with the creation of the first surveillance site in South Africa dating back to 1940. HDSS operate in a pre-specified geographical area, following prospectively a cohort of the population with the aim of monitoring information on key vital events (Chandramohan, 2008). The growing global presence of HDSS sites resulted in the creation of the International Network for the Demographic Evaluation of Populations and Their Health in Developing Countries (INDEPTH) in 1998, with a mandate of improving evidence in health based on a collaborative approach (INDEPTH Network, 2019). As of now, INDEPTH runs a network of 49 sites across 20 countries worldwide. Despite the great value of INDEPTH network, its current utilization is only complementary to the main source of key health indicators across all countries in Africa, i.e. the Demographic and Health Surveys (DHS) (Ye, 2012). Established in 1984, DHS are open-access, cross-sectional, national-representative, household-based surveys which are typically conducted every five years (The DHS program, 2019a). The nature of the collected data, together with the questionnaires and tests conducted by the program, have been considerably changed over the years by the continuous addition of indicators collected and the adoption of biomarker tests since 1995 (Corsi, 2012; The DHS program, 2019b). National-agencies, of the countries within which the surveys are conducted, are an integral part of the DHS project, which is funded by the United States Agency for International Development (USAID) (USAID, 2018) and implemented by the Inner City Fund (ICF) (ICF, 2019). Key components, relevant to our work, are measured indicators concerning under-five mortality, health interventions, such as malaria bed-nets and vaccination, maternal and household characteristics, as well as results from malaria parasitaemia and anemia haemoglobin testing. Complementary to the above, remote sensing and other open access data sources provide key information on the environmental and climatic conditions across SSA and therefore these data can be integrated to the DHS in order to reflect on the environmental and climatic conditions that each surveyed household was exposed. Examples of such data

sources include the U.S. Geological Survey-Earth Resources Observation Systems (USGSS) Data Portal, the Shuttle Radar Topography Mission (SRTM) and the Moderate Resolution Imaging Spectroradiometer (MODIS).

1.6 Rationale

The Sustainable Development Goals and the urgent need for the African countries to accelerate efforts on reducing U5M have inspired our work. Specifically our efforts address open research questions in the literature and contribute to the existing knowledge of the dynamics and main drivers of U5M in SSA. In terms of malaria-mortality modelling, previous studies utilizing survey data faced challenges associating malaria parasitaemia with U5M. Crucially, most of the modelling studies estimating malaria-related deaths have overlooked indirect causes of malaria-mortality, such as anemia. Both previous points raised are of great importance, as they might have resulted on underestimation of the malaria burden in SSA. In terms of health interventions and U5M, strong local-scale variation of the health interventions coverage and of the U5MR have been reported. Therefore, sub-national health inequalities and inequities might be present across various regions in SSA. Yet, there is sparsity of modelling studies identifying health inequities based on the effectiveness of health interventions on U5M. Further, another key issue is the contribution of the leading childhood diseases on fever burden. A common limitation shared among most studies investigating the contribution of illnesses to febrile response burden is the utilization of hospital-related data across a well-defined geographical area, i.e. specific, small in number, sites within a given country. Another shortage in the existing literature is that, to our knowledge, there are only a few modelling studies having tried to associate the contribution of diseases to fever at national level, but also exploring the geographical variation of the association within countries. Crucially, understanding the contribution of diseases on fever can assist with another widely reported issue, i.e. the over-diagnosis and hence over-treatment of malaria among under-five children with fever in SSA, as under-five children presenting fever in a health facility are mostly treated with antimalarial drugs. Furthermore, despite the proven

association between health interventions and under-five mortality in the African setting, there is a scarcity of studies taking a holistic approach and associating a large set of well proven for their effectiveness health interventions with under-five mortality, in order to take into account the synergistic effect of the scaling-up of multiple interventions. In addition, to our knowledge, there are no studies yet that have explicitly modelled the impact of interventions on the changes in mortality risk in Africa. Finally, given the recent attention attracted to Bayesian variable selection methods for spatially correlated data, there is a need for further assessing the sensitivity of such methods.

1.7 Thesis objectives and structure

The overall goal of this PhD thesis is to estimate spatio-temporal changes of U5M and of their association with health interventions and leading childhood diseases. To achieve that, we compiled and modeled approximately one million, cross-sectional, individual level data on under-five mortality and associated them with the malaria-anemia comorbidity prevalence and preventive or curative health interventions in sub-Saharan Africa. The core tool of our analysis was the development, implementation and interpretation of Bayesian geostatistical models with spatially varying coefficients. The point-by-point aims of this thesis are as follows:

1. To estimate the association of malaria parasitaemia, anemia, and malaria-anemia comorbidity with all-cause U5M and evaluate the potential of malaria-anemia comorbidity prevalence as a measure to quantify the malaria-related deaths in SSA (chapter 2);
2. To identify health inequities within countries in SSA due to the geographical variation in the association between curative or preventive health interventions and under-five mortality (chapter 3);
3. To assess the contribution of the leading causes of U5M, e.g. malaria parasitaemia, ARI and diarrhoea, in SSA on febrile response at national and sub-national scale. (chapter 4);

4. To estimate the association between health interventions and spatio-temporal changes of U5M in SSA. (chapter 5);
5. To evaluate and further develop geostatistical variable selection methods for models with spatially varying coefficients (chapter 6).

Chapter 2

Malaria-anemia comorbidity prevalence as a measure of malaria-related deaths in sub-Saharan Africa.

Isidoros Papaioannou[1,2], Jürg Utzinger[1,2], Penelope Vounatsou[1,2]

Affiliations:

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

This paper has been published in *Scientific Reports* 2019; **9:** 11323.

Abstract

Different methods and data sources have been utilized to determine the relationship between malaria and mortality in endemic countries. Most of these efforts have focused on deaths directly attributed to malaria, while they overlooked causes of mortality that might be indirectly related to the disease, for instance anemia. We estimated the association of malaria parasitaemia, anemia, and malaria-anemia comorbidity with all-cause under-five mortality and evaluated the potential of malaria-anemia comorbidity prevalence to quantify malaria-related deaths in sub-Saharan Africa. We analysed data from Demographic and Health Surveys (DHS) and employed Bayesian geostatistical models. Mortality hazard obtained from malaria-anemia comorbidity prevalence was up to 3·5 times higher compared to the hazard related to *Plasmodium* parasitaemia only. Malaria parasite prevalence alone could not always capture a statistically important association with under-five mortality. Geographical variation of the malaria-anemia comorbidity effect was observed in most, but not all, countries. We concluded that the malaria burden in sub-Saharan Africa is considerably underestimated when anemia in not taken into account and that the malaria-anemia comorbidity prevalence provides a useful measure of the malaria-related deaths.

2.1 Introduction

In 2016, malaria caused an estimated 216 million clinical episodes and 445,000 deaths (WHO, 2017a). Sub-Saharan Africa is the most afflicted region globally; indeed, 91% of all malaria deaths are concentrated in this part of the world. Children under the age of 5 years are at highest risk and might contribute up to 70% of all malaria-related deaths. Large-scale control efforts led to an 18% reduction in the incidence rate of malaria between 2000 and 2016 and six countries were certified by the World Health Organization (WHO) as malaria-free (WHO, 2017a). In 2016, the financial investment in malaria control and elimination was estimated at US$ 2·7 billion. Eliminating malaria by 2030 is part of the aspirations of the Sustainable Development Goals (SDGs) (UN, 2018). The 10 sub-Saharan African countries with the highest malaria burden experienced an increase in malaria incidence from 2015 to 2016, therefore the WHO recently launched the 10+1 initiative (WHO, 2018d) which aims to renew the focus on achieving the objectives set up by the Global Technical Strategy and Sustainable Development Goals.

Understanding the relation between malaria transmission and malaria-related mortality provides a useful insight in monitoring and evaluation of malaria control and elimination efforts. Furthermore, it helps in the planning of interventions since regions that are expected to have high malaria deaths could have interventions in place, such as availability of rapid diagnostic tests (RDTs) and Artemisinin-based combination therapy (ACT) in health facilities and information campaigns to prevent progression to severe malaria and therefore death. Several approaches and different data sources have been utilized to quantify malaria-related deaths. For example, the ‘Malaria Transmission Intensity and Mortality Burden across Africa’ (MTIMBA) project collected geo-referenced entomological data, which were linked with all-cause and malaria-specific mortality data derived from health and demographic surveillance systems (HDSS). A geostatistical analysis (Rumisha, 2014) of the MTIMBA data from the Rufiji HDSS in Tanzania found no significant relation between neonate, infant, and 1- to 4-year-old child survival with malaria transmission, the latter measured by the entomological

inoculation rate (EIR). In contrast, a 5-year collection of monthly data from the Kisumu HDSS in Kenya revealed a strong association between malaria-mortality and malaria transmission, the latter determined by the slide positivity rate (SPR) (Khagayi, 2017). Malaria survey and mortality data from the Demographic and Health Surveys (DHS) linked to malaria incidence from the Malaria Atlas Project (MAP) suggested a 57% decrease of malaria-related mortality during 1990 – 2015 (Gething, 2016). However, analyses of DHS mortality data from Mali and Malawi with historical malaria survey data were not able to capture a statistically significant malaria-mortality relation (Gemperli, 2006; Kazembe, 2007). A systematic analysis (Murray, 2012) of malaria mortality data between 1980 and 2010 estimated 1.238 million malaria deaths globally in 2010.

Most of the studies estimating malaria-related deaths overlooked indirect causes of malaria mortality (WHO, 2017b). Indeed, one of the main outcomes of *Plasmodium* infection is anemia caused by the rupture of red blood cells as part of the complex life cycle (Kai, 2008). *P. falciparum* leads to severe malarial anemia, which might be responsible for around one third of malaria deaths (Haldar, 2009). Righetti and colleagues (Righetti, 2013) reported a significant negative association between *Plasmodium* infections and haemoglobin concentrations. Another study reported stronger associations of anemia with malaria than iron deficiency or other nutritional, infectious, and genetic contributors (Foote, 2013). Despite the proven association between anemia, malaria, and mortality, modelling studies quantifying the association between malaria-anemia comorbidity and under-five mortality are rather few.

We estimated the association of malaria parasitaemia, anemia, and malaria-anemia comorbidity with all-cause under-five mortality and evaluated the potential of malaria-anemia comorbidity prevalence to quantify malaria-related deaths in sub-Saharan Africa. We compiled available household-based data collected from DHS and Malaria Indicator Surveys (MIS) and employed a rigorous Bayesian geostatistical modeling framework. We estimated within-country variation of the association between comorbidity and under-5 mortality, using

spatially varying coefficient models. Our results are presented at high spatial resolution, including model-based risk maps of malaria, anemia, and malaria-anemia comorbidity.

2.2 Methods

Data sources

We analysed household representative survey data from 16 sub-Saharan African countries that had available malaria, anemia, and mortality data at survey locations. The surveys were conducted between 2010 and 2017. A summary of the data extracted from the DHS is provided in Table 2.1.

For each child, the mortality-related data consisted of the age and vital status (i.e. alive or dead) of the child. Furthermore, we considered in the analysis information on maternal, household and individual child characteristics. Maternal data included the educational attainment and literacy of the mother, age at pregnancy, delivery methods, other pregnancy terminations, and time intervals succeeding or preceding birth. Each child was classified into a socioeconomic category, using a household-based asset index which was already available in the data (Rutstein, 2004). We also obtained information on the living standards of each child (e. g. improved sanitation facilities, improved drinking water sources, and open defecation practises). Individual child covariates included the sex of a child, birth order, breastfeeding practises of the mothers, and vaccination status.

Malaria and anemia data were extracted from readily available MIS. Some of the surveys used a combination of malaria microscopy and rapid diagnostic tests (RDTs) to diagnose malaria, while others used RDTs only. We classified the anemia level of a child based on guidelines put forward by the World Health Organization (WHO), as follows: (i) moderate or severe anemic if haemoglobin (Hb) levels were below 100 g/l and (ii) severe anemic if Hb levels were below 70 g/l. We also considered indicators of malaria intervention coverage at cluster level such as the proportion of households reporting indoor residual spraying (IRS), Artemisinin-based combination therapy coverage, the proportion of use or ownership of insecticide-treated nets (ITN), and the prevalence of iron supplementation in order to adjust

for interventions. We defined comorbidity as malaria parasitaemia with moderate or severe anemia and severe comorbidity as parasitaemia with severe anemia. The DHS Program maintains strict standards for protecting the privacy of respondents and household members in all DHS surveys. Before each interview or biomarker test is conducted, an informed consent statement is read to the respondent, who may accept or decline to participate. Also, verbal informed consent for each parasitaemia test is provided by the child's parent/guardian/caregiver on behalf of children less than 5 years before the test is conducted. Verbal consent is conducted by the interviewer reading a prescribed statement to the respondent and recording in the questionnaire whether or not the respondent consented or assent is provided. The interviewer signs his or her name attesting to the fact that he/she read the consent statement to the respondent. Written consent is not included (https://dhsprogram.com/publications/publication-dhsm7-dhs-questionnaires-and-manuals.cfm).

We extracted environmental and climatic data from remote sensing and other open access data sources. We downloaded from Moderate Resolution Imaging Spectroradiometer (MODIS) the normalized difference vegetation index (NDVI) and the land surface temperature (LST) at 1x1 km^2 spatial resolution. We also obtained from MODIS the land cover type (LC) and distance from nearest water bodies (DWATER) at 0·5x0·5 km^2 spatial resolution. LC was recoded to the following categories: forests, grasslands, and croplands. Rainfall data were obtained from the U.S. Geological Survey-Earth Resources Observation Systems (USGSS) Data Portal, while altitude data were extracted from the Shuttle Radar Topography Mission (SRTM) at 0·5x0·5 km^2 spatial resolution. Locations were classified to rural or urban residence according to the Global Rural and Urban Mapping Project.

Statistical analysis

We fitted separate Bayesian geostatistical Weibull survival models, for each one of the morbidity indicators considered in our study, i.e. malaria parasitaemia, moderate/severe anemia, severe anemia, moderate comorbidity and severe comorbidity, in order to assess

their effect on all-cause under-five mortality. The outcome of the survival model outcome the age of death or current age of a child (in months), together with the corresponding censoring status, i.e. alive children were considered as censored observations. The indicators measured cluster-level prevalence and thus each child was associated with the prevalence at the corresponding cluster. Therefore, the prevalence of malaria parasitaemia, anemia, and malaria-anemia comorbidity at a given location was treated as an exposure that children receive at that location. For each model, the association between the mortality outcome and the corresponding indicator was adjusted for confounders, which were selected based on rigorous variable selection procedures, carried out separately for each country. In particular, we fitted separate Weibull survival geostatistical models with the candidate covariates and included only the statistically important ones (i.e. Bayesian credible interval did not include zero). We used throughout the paper terminology consistent with the Bayesian inference and call statistically important effects regression coefficients that are known as statistically significant in frequentist inference. The variance inflation factor was used to exclude the highly correlated predictors (O'Brien, 2007). The final models included different sets of covariates for each country in order to estimate the malaria parasitaemia, anemia, and comorbidity effect from the most parsimonious model. We included spatially random effects at cluster level modelled by a Gaussian process with Matérn covariance matrix (Banerjee, 2014).

An additional analysis was carried out to assess the sensitivity of the results to the selected confounders. Specifically, we re-fitted the parasitaemia and comorbidity prevalence models using a common set of confounders across all countries based on the biggest set derived by combining all country-specific confounders. Furthermore, we re-fitted the models with the comorbidity prevalence adjusting for parasitaemia prevalence only (prevalence of children with positive parasite test which are not moderate or severe anemic) and moderate-to-severe anemia prevalence only (prevalence of moderate or severe anemic children which are not positive for a malaria parasite test).

We identified the geographical distribution in the effects of moderate comorbidity on all-cause under-five mortality, by fitting a spatially varying coefficient model for each country, assuming a spatially continuous Gaussian process on the comorbidity effect. Estimates were summarised by the posterior median, 2·5 and 97·5 quantiles. The models were adjusted for malaria or anemia interventions, climatic, maternal, individual child, and household characteristics.

We used geostatistical models to estimate disease risk surfaces and created national maps based on pixel-based estimates. Environmental predictors and spatial random effects were modelled on the logit scale. Prediction was carried out into a grid formed by 2*x*2 km^2 resolution pixels.

The analysis was conducted in R software (version 3.3.2). We used the integrated nested Laplace approximations (INLA) approach (Rue, 2009) to perform fast approximate Bayesian inference. Model details are provided in the Appendix.

2.3 Results

A total of 8,116 unique locations from 16 sub-Saharan Africa countries were included in this study. Of these, Angola, Benin, and Uganda have the highest number of locations. The earliest survey from which we extracted data is from Burkina Faso, conducted in 2010, while the latest is from Burundi, conducted in 2016-2017. Malaria prevalence among children under 5 years ranged from 2.2% in Rwanda to 76·1% in Burkina Faso. All countries of West Africa, except Senegal, had higher malaria prevalence compared to East African countries. The highest moderate/severe anemia prevalence was found in Burkina Faso (69.6%). Burkina Faso, Guinea, and Cameroon showed, in descending order, the highest under-five mortality rate. Despite relatively high malaria parasitaemia and anemia levels, Ghana revealed the second lowest under-five mortality rate among the 16 countries examined. Burkina Faso, Burundi, Mozambique, and Uganda were the four countries characterized by a higher malaria burden compared to the anemia burden. All countries except Burkina Faso (11.1%) had severe anemia levels below 10% with the lowest rate observed in Rwanda

(0·7%). Overall, the mean malaria prevalence was 31·1%, mean moderate/severe anemia was 39·7%, and mean severe anemia was 3·7%.

Our estimates (Fig. 2.1) highlight a statistically important effect of comorbidity on under-five mortality across all countries, except Rwanda and Tanzania. We identified a statistically important effect of malaria-anemia comorbidity on under-five mortality in 14 out of 16 countries (except Rwanda and Tanzania), while malaria prevalence alone was statistically important in half of the countries. In Benin, the hazard ratio (HR) of comorbidity was 3·5 times the hazard ratio of parasitaemia alone. The HR of comorbidity was approximately 50% higher in Cameroon, Mozambique, and Togo and approximately 30% higher in Uganda and the Democratic Republic of the Congo compared to the HR of parasitaemia alone. We observed the lowest difference in hazard ratios in Angola and Mali, with 7% and 3% higher malaria-anemia comorbidity hazard, respectively. Burkina Faso was the only country with an important comorbidity effect and no statistically important difference in the hazard ratios of comorbidity and parasitaemia.

Using the same set of confounders across all countries, we found a statistically important association of malaria-anemia comorbidity prevalence with under-five mortality in 9 out of 16 countries examined (Appendix, Table 2.4). Similarly to the results above, the comorbidity prevalence was associated with mortality in countries which did not have an important parasitaemia-mortality relation. In countries with a significant parasitaemia-mortality association, the coefficient of the comorbidity was greater than that of parasitaemia (except Burkina Faso, as observed with country-specific covariates). Estimates for the comorbidity prevalence, adjusted for parasitaemia only prevalence, anemia only prevalence and country-specific confounders, were higher in 7 out of 11 countries compared to unadjusted models for parasitaemia and anemia only prevalences (Appendix, Table 2.5). However, in the adjusted analyses the comorbidity was not statistically important in Angola, Guinea and Mali. Severe disease expressed either as severe anemia or severe comorbidity has the highest burden on under-five mortality. As all countries have relatively low severe anemia prevalence, there is high uncertainty in our estimates, contributing to important associations

only in 4 out of the 16 countries investigated (Fig. 2.1). However, when the severe comorbidity-mortality relation is captured, the HR can be more than five times higher to that of comorbidity, as observed in Cameroon and Côte d'Ivoire. Similarly, when we identify an important severe anemia-mortality association, the HR estimates of severe anemia are greater than the ones of comorbidity, albeit lower than that of severe comorbidity. Hence it can be inferred that severe disease has the greatest impact on under-five survival.

Figure 2.2 demonstrates the geographical distribution of the comorbidity association with all-cause under-five mortality. All countries with a statistically important comorbidity effect at national level demonstrate subnational geographical variations. In some countries (i.e. Benin, Côte d'Ivoire, Democratic Republic of the Congo and Mozambique) there are large areas with a statistically important comorbidity effect. In others, (i.e. Angola, Burkina Faso, Burundi, Cameroon and Ghana) the effect is rather focal.

Model-based risk surfaces of the comorbidity effect on under-five mortality are presented in Figure 2.3 and of malaria parasitaemia and anemia in the Appendix. In some countries, we observed strong spatial patterns of moderate-to-severe anemia risk surfaces while in others there are only a few high risk hotspots. Noteworthy are areas in which coincide high malaria parasitaemia risk coincides with densely populated surfaces, for example in northern Burundi, central-south Cameroon, central Guinea, and north-east Mozambique. Severe anemia-malaria maps refine all areas of high risk within a country and demonstrate regions of severe disease, i.e. areas between Lunda Sul and the Moxico border in Angola, Ngozi province in Burundi, Mopti province in Mali, the south-eastern part of Senegal, North Tanzania, and the south-western part of Côte d'Ivoire.

2.4 Discussion

This is the first study assessing the association of malaria-anemia comorbidity with child mortality using routinely collected household survey data from DHS and MIS. We employed Bayesian geostatistical modeling and rigorous variable selection procedures to assess the geographical variation of the comorbidity effect on under-five mortality. We included in the

analysis survey data with information on mortality, anemia and malaria at the same location to avoid potential bias from surveys misaligned in space and time.

Our findings demonstrate a strong association between under-five mortality and comorbidity burden, suggesting that higher prevalence of the comorbidity are related to higher mortality rates. Low malaria parasite prevalence, low under-five mortality, and moderate-to-high malaria intervention coverage were observed in countries with no statistically important comorbidity effects (i.e. Rwanda and Tanzania). In countries with a statistically important malaria-mortality relation, with the exception of Burkina Faso, the mortality hazard of comorbidity was higher than the mortality hazard of parasitaemia. A cross-sectional study in Rwanda found that children with moderate/severe anemia had a four times higher risk of malaria infection (Kateera, 2015). A prospective observational study in Gabon (Bouyou-Akotet, 2009) showed that moderate/severe malarial anemia was strongly associated with parasitaemia in children aged below four years. Other studies (Osterbauer, 2012; Magalhães, 2013; Ceesay, 2015; Moraleda, 2017; Menon, 2015) also suggested a strong relationship between malaria and anemia, irrespective of the malaria prevalence levels (McCuskee, 2014). Our work demonstrates that moderate anemic children with malaria parasites have higher mortality hazards compared to those children with only parasitaemia. When a common set of confounders was used across all models, we observed an important comorbidity-mortality association in fewer countries, albeit these countries did not have an important parasitaemia-mortality association. This modelling approach incorporates, in each model, unnecessary covariates and it does not take into account multicollinearity which can bias parameter estimation. The mortality-comorbidity association was not statistically important in Angola, Guinea and Mali when the model was adjusted for parasitaemia only and anemia only prevalences. In fact these countries have the highest anemia only prevalence (together with Senegal and Tanzania) which may explain the variation in mortality. We have adjusted for the malaria/anemia only to avoid high correlation with the comorbidity prevalence.

Among the factors investigated in our study, severe comorbidity has the highest impact on all-cause under-five mortality. Due to the low severe comorbidity prevalence, we obtained higher uncertainty in our estimates and thus we were only able to find an important association with under-five mortality in Cameroon, Côte d'Ivoire, Democratic Republic of the Congo, and Uganda. In these countries, the magnitude of association of severe comorbidity is far greater compared to that of other conditions and types of diseases examined. A study from the University Teaching Hospital in Ebonyi state in Nigeria showed that malaria accounted for approximately 77% of all children with severe anemia admission in the hospital (Muoneke, 2012) and associated severe anemia with high under-five mortality rates. In western Kenya, an association of 85% between parasitaemia and severe anemia in hospital admissions is reported (Obonyo, 2007), with severe anemia contributing to approximately half of malaria-related deaths. Severe comorbidity accounts for more than 30% in paediatric mortality (Perkins, 2011). Apart from profound results on children survival, severe malarial anemia could also be associated with long-term impairment in cognitive ability (Bangirana, 2014).

Our findings revealed considerable geographical variation of the malaria-anemia comorbidity effect on all-cause under-five mortality in most countries. Several geographical areas with important comorbidity association coincide with known areas in which high mortality and low antimalarial treatment or ITN coverage exists (Gething, 2016), for instance areas in Lunda Norte, Huila and Malanje provinces in Angola, Orientale, Katanga and Kivu provinces in Democratic Republic of the Congo, parts in north Cameroon and central Côte d'Ivoire, Faranan and Kankan provinces in Guinea and most of the areas shown in Mali. Additionally, some of the areas reside in provinces that experienced severe malaria burden at the year that the DHS-MIS was conducted, for instance the Orientale province in Democratic Republic of the Congo (DOCTORS WITHOUT BORDERS, 2018).

In countries lacking a clear malaria-mortality relation, we found an important anemia-mortality association. Democratic Republic of the Congo, Mozambique, and Uganda are the only countries in this study showing that both malaria parasite and anemia prevalence are

related to mortality. Greenwell and Neuman (Greenwell, 2018) supported that anemia is an important indicator for monitoring malaria burden. Based on our findings, the malaria-anemia comorbidity indicator could be used to monitor severe disease.

Our work assessed the potential of anemia and malaria parasitaemia comorbidity prevalence to estimate malaria-related under-five mortality. Our findings confirm earlier results showing that malaria parasite prevalence alone fails to fully capture a statistically important association with mortality. Presence of malaria parasites is not directly linked to severe malaria that leads to death and thus statistical models may fail to capture the known association between malaria and mortality. In contrast, the comorbidity indicator can identify cases that experience severe disease, and hence, it is better related to mortality. It follows that the effect of comorbidity prevalence on all-cause under-five mortality can be a better indicator of malaria-related deaths than malaria prevalence alone, in survey data. Studies should further evaluate the use of comorbidity indicator in estimating malaria deaths for under-five children using malaria-specific mortality data.

Our work does not take into account changes in prevalence that might have been occurred over the last five years that mortality has been studied. This assumption may influence data stemming from the most recent surveys, when the scaling up of malaria interventions had already taken place. Yet, our analysis includes mostly surveys conducted before 2013, thus before the further scaling up of interventions. In our study the estimated effects are quantifying associations and by no means imply causal relations. Children with severe malaria anemia are less likely to participate in DHS or MIS, leading to underestimation of this effect.

Acknowledgements

The authors are grateful to the DHS for making the data available and all collaborating institutions whose contributions have enabled this study. This work was funded by the European Research Council (ERC) Advanced Grant project no. 323180.

Ethical approval and consent to participate

In this study we analysed secondary data made available by the Demographic Health Survey (DHS) MEASURE. According to survey protocols and related documents of the surveys, ethical approval was obtained from the Institutional Review Board of International Consulting Firm (ICF) of Calverton, Maryland, USA, and also from the national ethical committees in the countries that the surveys were contacted. Details of ethical clearance are published in the DHS reports available at https://dhsprogram.com/publications/index.cfm.

Competing interests

We declare no competing interests.

Role of funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the paper. The authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Contributors

IP processed and analysed the data, interpreted the results, and wrote the manuscript. PV conceptualised the project and assisted in statistical analysis. JU and PV revised the manuscript and provided important intellectual content. All authors approved the final version of the manuscript.

Availability of Materials and Data

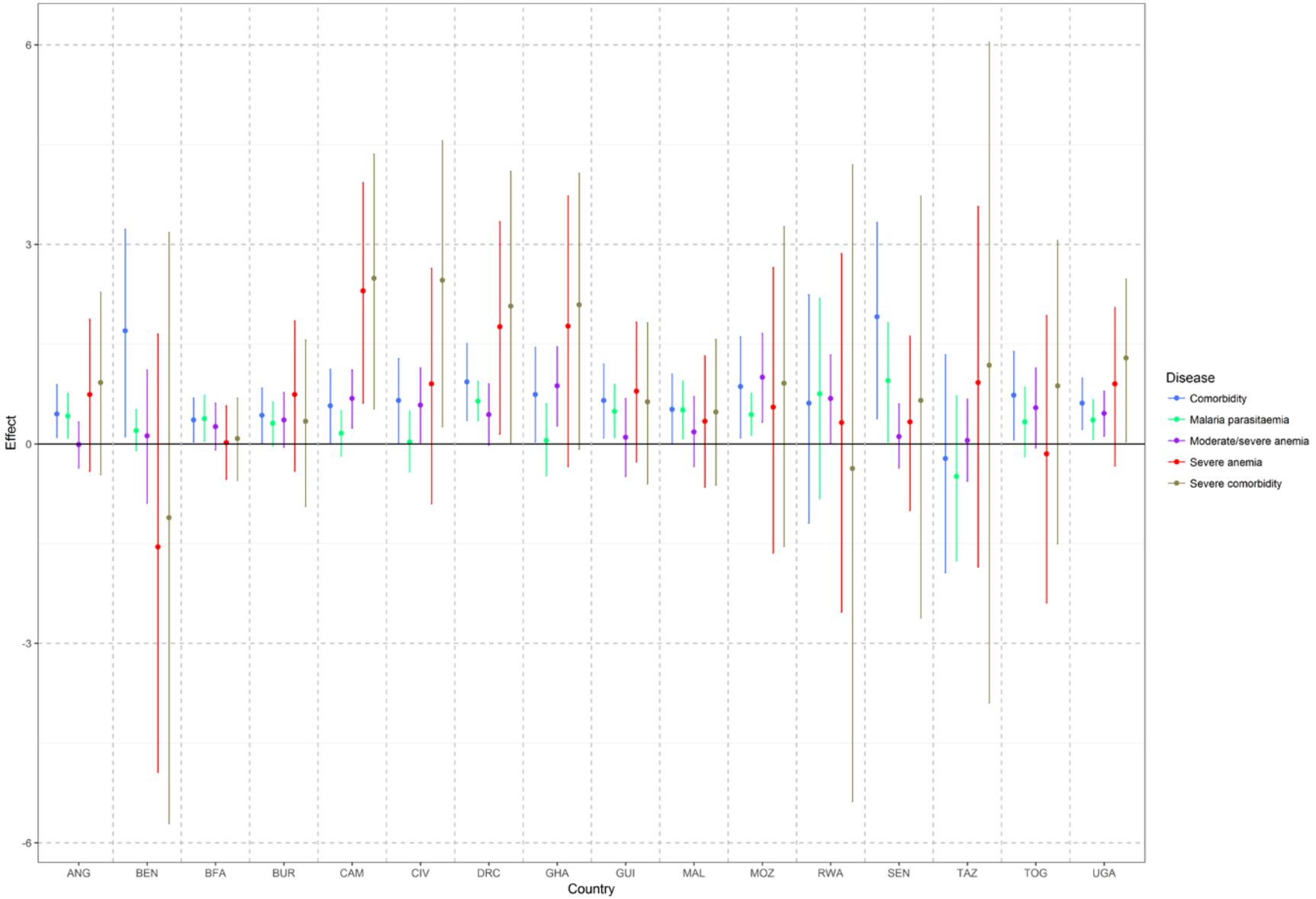
The study data are available upon request from the Demographic and Health Surveys program (https://dhsprogram.com/).

Table 2.1: Descriptive analysis of malaria survey data in 16 sub-Sahara African countries, collected from the DHS programme.

Country	Survey period	Number of locations	Malaria parasitaemia* (%)	Moderate/ severe anemia* (<100 g/l)	Severe Anemia* (<70 g/l)	Moderate comorbidity*	Severe comorbidity*	ITN**	U5M***
Angola	10/2015 - 3/2016	625	13·1% (R)	34·0	2·2	10·1	1·5	31·0	68
Benin	12/2011 - 4/2012	747	28·4% (M)	32·1	3·0	10·9	1·0	79·8	70
Burkina Faso	1/2010 - 12/2010	542	76·1% (R)	69·6	11·1	53·2	8·5	56·9	129
Burundi	10/2016 - 3/2017	552	37·9% (R)	36·3	3·6	21·0	2·9	46·1	78
Cameroon	1/2011 - 8/2011	577	30·0% (R)	32·8	1·7	16·8	1·4	36·4	122
Côte d'Ivoire	12/2011 - 5/2012	341	41·5% (R)	49·7	3·3	27·7	2·4	67·3	108
DRC	8/2013 - 2/2014	492	22·6% (M)	27·1	2·3	11·9	1·4	70·0	104
Ghana	9/2014 - 12/2014	423	36·4% (R)	39·1	2·2	25·0	2·1	68·3	60
Guinea	6/2012 - 10/2012	300	43·9% (M)	52·4	7·6	27·0	4·9	47·0	123
Mali	11/2012 - 2/2013	413	51·6% (R)	60·8	9·3	33·4	6·7	84·4	95
Mozambique	5/2011 - 12/2011	610	38·3% (R)	26·6	1·4	21·7	2·7	51·4	97
Rwanda	11/2014 - 4/2015	492	2·2% (M)	15·7	0·7	1·5	0·2	80·6	50
Senegal	10/2010 - 4/2011	385	2·9% (M)	53·2	4·9	2·5	0·6	62·9	72
Tanzania	8/2015 - 2/2016	608	5·6% (M)	31·2	1·6	2·4	0·2	65·6	67
Togo	10/2013 - 4/2014	330	36·4% (M)	44·8	2·4	22·5	1·8	65·4	88
Uganda	6/2016 - 12/2016	679	30·3% (R)	29·2	2·3	17·1	1·8	78·4	64

* Representing prevalence; ** Insecticide-treated net ownership; *** Under-five mortality (per 1,000); **** M for microscopy rest; R for RDT.

Figure 2.1: Bayesian estimates (posterior median, 95% BCI) of malaria parasitaemia, moderate/severe anemia (<100 g/l), severe anemia (<70 g/l), comorbidity, and severe comorbidity prevalence on under-five mortality in 16 sub-Sahara African countries. Models were adjusted for climate, maternal and household characteristics, malaria- anemia interventions, and individual child covariates.



ANG, Angola; BEN, Benin; BFA, Burkina Faso; BUR, Burundi; CAM, Cameroon; CIV, Côte d'Ivoire; DRC, Democratic Republic of the Congo; GHA, Ghana; GUI, Guinea; MAL, Mali; MOZ, Mozambique; RWA, Rwanda; SEN, Senegal; TAZ, Tanzania; TOG, Togo; UGA, Uganda

Figure 2.2: Posterior median of the malaria-anemia comorbidity effect on all-cause under-five mortality. Effects which are not statistically important (pixel-level posterior distribution includes zero) are indicated in grey colour.

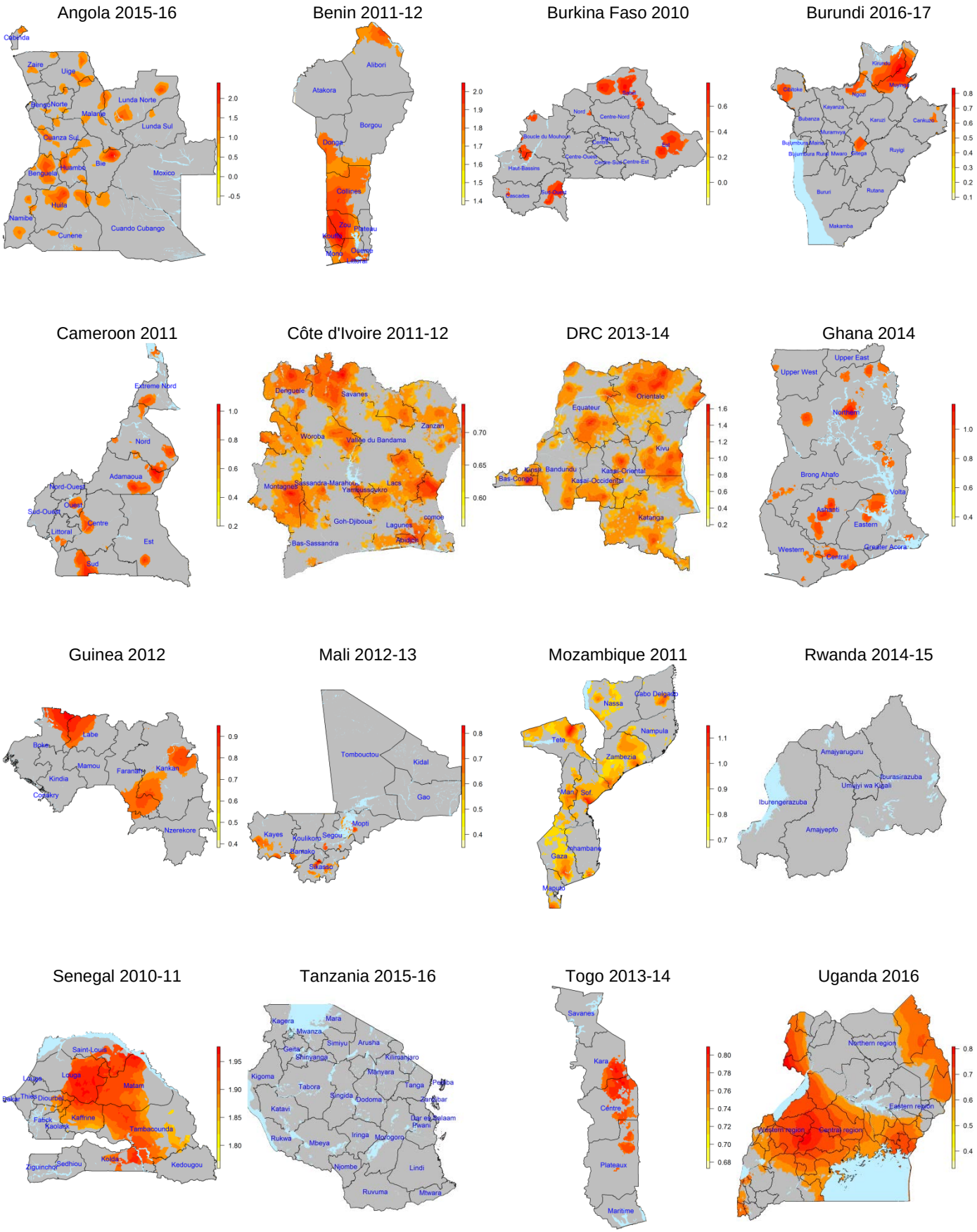
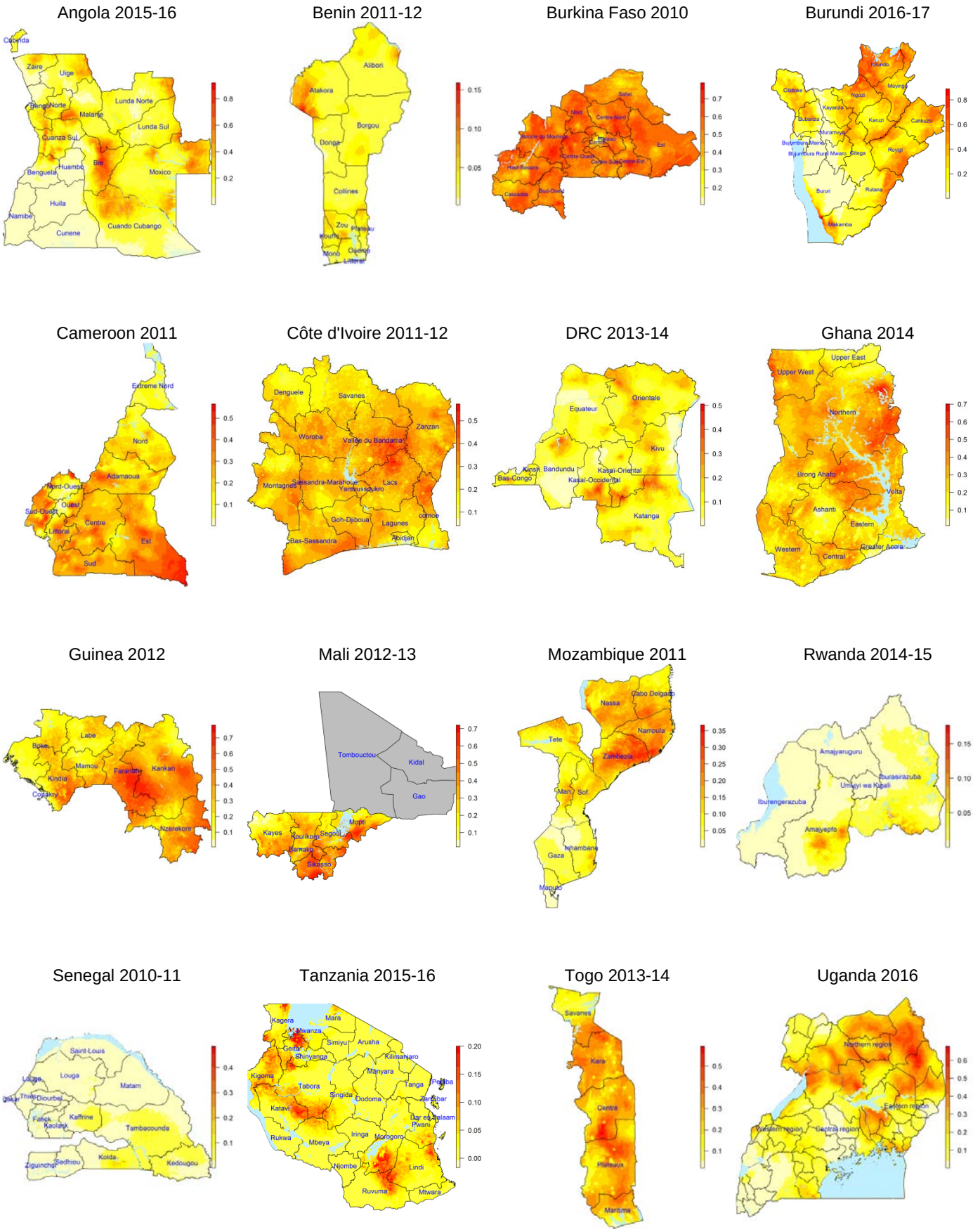


Figure 2.3: Posterior median of malaria parasitaemia and moderate or severe anemia comorbidity risk estimates at 2x2 km^2 from Bayesian geostatistical models for 16 sub-Sahara African countries.



2.5 Appendix

Model specification

Bayesian geostatistical Weibull survival model

We developed a geostatistical Weibull survival model to assess the association of malaria, anemia, and malaria-anemia comorbidity with under-five mortality in 16 sub-Saharan Africa countries. Let $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$ be the set of surveyed locations with observed malaria, anemia, and mortality data. We defined T_{ji} the age of death for the j-th child at location s_i and X_{ji} the row-vector of the associated covariates. We assumed a Weibull lifetime distribution for time T_{ji} with shape parameter α and scale parameter λ and probability density function $f(t_{ji}) = a\ {t_{ji}}^{a-1}\ \lambda \exp(-\lambda\ {t_{ji}}^{a}), \alpha > 0, \lambda > 0$. Under this specification, the corresponding hazard $h(t)$ and survival $S(t)$ have the following form: $h(t_{ji}) = a\ {t_{ji}}^{a-1}\ \lambda$ and $S(t_{ji}) = \exp(-\lambda {t_{ji}}^{\alpha})$. We introduced predictors on the scale parameter using a log link function, that is: $\lambda_{ji} = \exp(\eta_{ji})$, where η_{ji} is the linear predictor defined by the equation,

$$\eta_{ji} = \beta_0 + \sum_{\kappa=1}^{K} \beta_\kappa x_{ji\kappa} + \xi_i,$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_\kappa)^T$ is the vector of K regression coefficients, $x_{ji\kappa}$ represents the value of the κ-th predictor and $\xi(s) = (\xi_1, \xi_2, \dots, \xi_n)^\tau$ are multivariate Normal distributed random variables that take into account for spatial correlation in the response among the s locations. The multivariate Normal distribution followed a zero-mean specification with covariance between locations s_1 and s_2 described by the Matérn function, i.e. $\xi(s) \sim N(0, \Sigma_1)$ and $\Sigma_1(s_1, s_2) = \frac{\sigma_1^2 (\kappa_1\ d(s_1,s_2))^{\nu} K_\nu(\kappa_1\ d(s_1,s_2))}{\Gamma(\nu)\ 2^{\nu-1}}$, with σ_1^2 being the spatial process variance, $d(s_1, s_2)$ the distance between locations s_1 and s_2 and κ_1 the scaling parameter. K_ν is the modified Bessel function of second kind and order ν. Spatial range (r_1) describes the distance at which the spatial correlation becomes negligible and under the Matérn specification is defined as $r_1 = \frac{\sqrt{8}}{\kappa_1}$. We selected normally distributed priors $N(0{,}10)$ for

regression coefficients and the default priors of INLA for the spatial hyperparameters and the shape of the Weibull distribution.

Spatially varying coefficient modelled via a spatially continuous Gaussian process

The above model was extended to include a spatially varying comorbidity effect. In particular, the linear predictor was written as follows:

$$\eta_{ji} = \beta_0 + b_i x_{ji1} + \sum_{\kappa=2}^{K} \beta_\kappa x_{ji\kappa} + \xi_i,$$

where b_i represents a spatially continuous Gaussian process, $\boldsymbol{b} \sim N(\beta_1, \Sigma_2)$ and $\boldsymbol{b} = (b_1, \dots, b_n)^\tau$. The mean β_1 represents the global comorbidity effect (i.e. at national scale). Σ_2 is a Matérn covariance function with spatial process variance σ_2^2 and scaling parameter κ_2 (section 1.1.1). The geographical distribution of the comorbidity effect was estimated by predicting the Gaussian process over a gridded surface of 2x2 km^2 spatial resolution.

Bayesian geostatistical binomial model

We developed Bayesian geostatistical binomial models to estimate disease risk surfaces, i.e. malaria parasitaemia, moderate/severe anemia, severe anemia, comorbidity, and severe comorbidity. We defined Y_i as the number of diseased under-five children at location s_i, $s = (s_1, s_2, \dots, s_n)^\tau$, $s_i \subset R^2$ and N_i the total number of children observed at that location. We assumed that for each location s_i the number of positive to the disease children Y_i follows a binomial distribution, i.e., $Y_i \mid N_i, \pi_i \sim Bin(N_i, \pi_i)$, with π_i being the disease risk at that location. The disease risk π_i was reparameterised in terms of predictor variables and regression parameters by using the logit link function as follows:

$$logit(\pi_i) = \beta^\tau X_i + \varphi_i,$$

where X_i is a set of environmental predictors observed at location s_i and $\beta = (\beta_0, \beta_1, \dots, \beta_k)^T$ the corresponding regression coefficients. The latent spatial process $\varphi(s) = (\varphi_1, \varphi_2, \dots, \varphi_n)^\tau$ follows the same specification as defined above.

Environmental and climatic data

Table 2.2: Sources of environmental and climatic data.

Data	Source	Spatial resolution
Annual average Normalised Difference Vegetation Index (NDVI)	MODIS	1x1 km^2
Annual average Day and Night Land Surface Temperature (LST)	MODIS	1x1 km^2
Land Cover Type (LC)	MODIS	0·5x0·5 km^2
Distance from water bodies (DWATER)	MODIS	0·5x0·5 km^2
Annual average Rainfall	USGSS	8x8 km^2
Altitude (Digital Elevation model)	SRTM	0·5x0·5 km^2
Urban rural extent	GRUMP	1x1 km^2

Additional Results

Table 2.3: Bayesian estimates (posterior median, 95% BCI) of malaria, anemia, and malaria-anemia comorbidity prevalence on under-five mortality. Models were adjusted for country-specific confounders related to climate, mother, individual child and household characteristics as well as malaria- anemia interventions.

Condition Country	Malaria parasitaemia	Moderate/ severe anemia	Severe anemia	Malaria-moderate/ severe anemia	Malaria/ severe anemia
Angola	**0·42 (0·09,0·76)**	-0·01 (-0·36,0·33)	0·74 (-0·41,1·87)	**0·45 (0·1,0·89)**	0·92 (-0·46,2·28)
Benin	0·20 (-0·10,0·52)	0·12 (-0·89,1·11)	-1·55 (-4·94,1·65)	**1·70 (0·11,3·23)**	-1·11 (-5·71,3·18)
Burkina Faso	**0·38 (0·04,0·73)**	0·26 (-0·09,0·61)	0·02 (-0·53,0·57)	**0·36 (0·03,0·69)**	0·08 (-0·55,0·69)
Burundi	0·31 (-0·03,0·63)	0·36 (-0·05,0·77)	0·74 (-0·41,1·85)	**0·43 (0·01,0·84)**	0·34 (-0·94,1·56)
Cameroon	0·16 (-0·18,0·50)	**0·68 (0·24,1·11)**	**2·30 (0·61,3·93)**	**0·57 (0·02,1·12)**	**2·49 (0·53,4·36)**
Côte d'Ivoire	0·03 (-0·42,0·49)	**0·58 (0·02,1·14)**	0·90 (-0·90,2·64)	**0·65 (0·02,1·28)**	**2·46 (0·26,4·56)**
DRC	**0·64 (0·35,0·94)**	0·44 (-0·02,0·90)	**1·76 (0·15,3·34)**	**0·93 (0·35,1·51)**	**2·07 (0,4·10)**
Ghana	0·05 (-0·48,0·60)	**0·87 (0·27,1·46)**	1·77 (-0·34,3·73)	**0·74 (0·03,1·45)**	2·09 (-0·08,4·07)
Guinea	**0·49 (0·10,0·89)**	0·10 (-0·49,0·68)	0·79 (-0·27,1·83)	**0·65 (0·09,1·20)**	0·63 (-0·60,1·82)
Mali	**0·51 (0·08,0·94)**	0·18 (-0·34,0·71)	0·34 (-0·65,1·32)	**0·52 (0·01,1·05)**	0·48 (-0·62,1·57)
Mozambique	**0·44 (0·13,0·76)**	**1·00 (0·33,1·66)**	0·55 (-1·64,2·65)	**0·86 (0·09,1·61)**	0·91 (-1·54,3·27)
Rwanda	0·75 (-0·82,2·19)	**0·68 (0·01,1·34)**	0·32 (-2·53,2·86)	0·61 (-1·19,2·24)	-0·37 (-5·38,4·20)
Senegal	**0·95 (0·03,1·82)**	0·11 (-0·36,0·60)	0·33 (-1·00,1·62)	**1·91 (0·38,3·33)**	0·65 (-2·62,3·73)
Tanzania	-0·49 (-1·76,0·72)	0·05 (-0·56,0·67)	0·92 (-1·85,3·57)	-0·22 (-1·94,1·34)	1·18 (-3·90,6·04)
Togo	0·33 (-0·19,0·85)	0·54 (-0·06,1·14)	-0·15 (-2·39,1·93)	**0·73 (0·06,1·39)**	0·87 (-1·50,3·06)
Uganda	**0·36 (0·07,0·66)**	**0·46 (0·12,0·79)**	0·90 (-0·33,2·05)	**0·61 (0·22,0·99)**	**1·29 (0·03,2·48)**

Table 2.4: Bayesian estimates (posterior median, 95% BCI) of malaria parasitaemia and malaria-anemia comorbidity prevalence on under-five mortality. Models were adjusted for a common set of confounders based on the biggest set derived by combining all country-specific confounders.

Condition Country	Malaria parasitaemia	Malaria-moderate/ severe anemia
Angola	0.23 (-0.11,0.57)	0·21 (-0·24,0·66)
Benin	0.20 (-0.10,0.50)	**1.70** **(0.15,3.19)**
Burkina Faso	**0.46** **(0.12,0.81)**	**0.42** **(0.09,0.75)**
Burundi	**0.34** **(0.02,0.66)**	**0.43** **(0.01,0.83)**
Cameroon	0.17 (-0.16,0.51)	0.48 (-0.07,1.02)
Côte d'Ivoire	0.24 (-0.18,0.68)	**0.87** **(0.26,1.47)**
DRC	**0.70** **(0.41,0.99)**	**1.04** **(0.47,1.60)**
Ghana	0.10 (-0.41,0.61)	0.64 (-0.02,1.29)
Guinea	**0.62** **(0.24,1.00)**	**0.91** **(0.39,1.43)**
Mali	0.32 (-0.08,0.71)	0.33 (-0.16,0.82)
Mozambique	0.42 (-0.20,1.04)	0.63 (-0.16,1.39)
Rwanda	0.93 (-0.63,2.32)	0.77 (-1.02,2.36)
Senegal	**1.14** **(0.25,1.98)**	**1.85** **(0.40,3.20)**
Tanzania	-0.65 (-1.91,0.53)	-0.54 (-2.24,1.01)
Togo	0.45 (-0.06,0.95)	**0.78** **(0.13,1.41)**
Uganda	0.29 (-0.01,0.57)	**0.55** **(0.16,0.93)**

Table 2.5: Bayesian estimates (posterior median, 95% BCI) of malaria-anemia comorbidity on under-five mortality. Models were adjusted for parasitaemia only prevalence (children without moderate/severe anemia), anaemia only prevalence (children without malaria parasitaemia) and country-specific confounders.

Condition Country	Malaria-moderate/ severe anemia
Angola	0.31 (-0.16,0.78)
Benin	**1.78** **(0.19,3.31)**
Burkina Faso	**0.64** **(0.12,1.18)**
Burundi	**0.44** **(0,0.87)**
Cameroon	**0.70** **(0.12,1.26)**
Côte d'Ivoire	**0.65** **(0.01,1.30)**
DRC	**0.75** **(0.14,1.35)**
Ghana	**0.85** **(0.11,1.57)**
Guinea	0.43 (-0.22,1.06)
Mali	0.42 (-0.17,1.00)
Mozambique	**1.03** **(0.24,1.80)**
Rwanda	0.50 (-1.30,2.12)
Senegal	**1.89** **(0.35,3.32)**
Tanzania	-0.11 (-1.83,1.46)
Togo	**0.72** **(0.01,1.42)**
Uganda	**0.64** **(0.24,1.03)**

Figure 2.4: Bayesian geostatistical model based malaria parasitaemia risk estimates in 16 sub-Sahara African countries at 2x2 km^2.

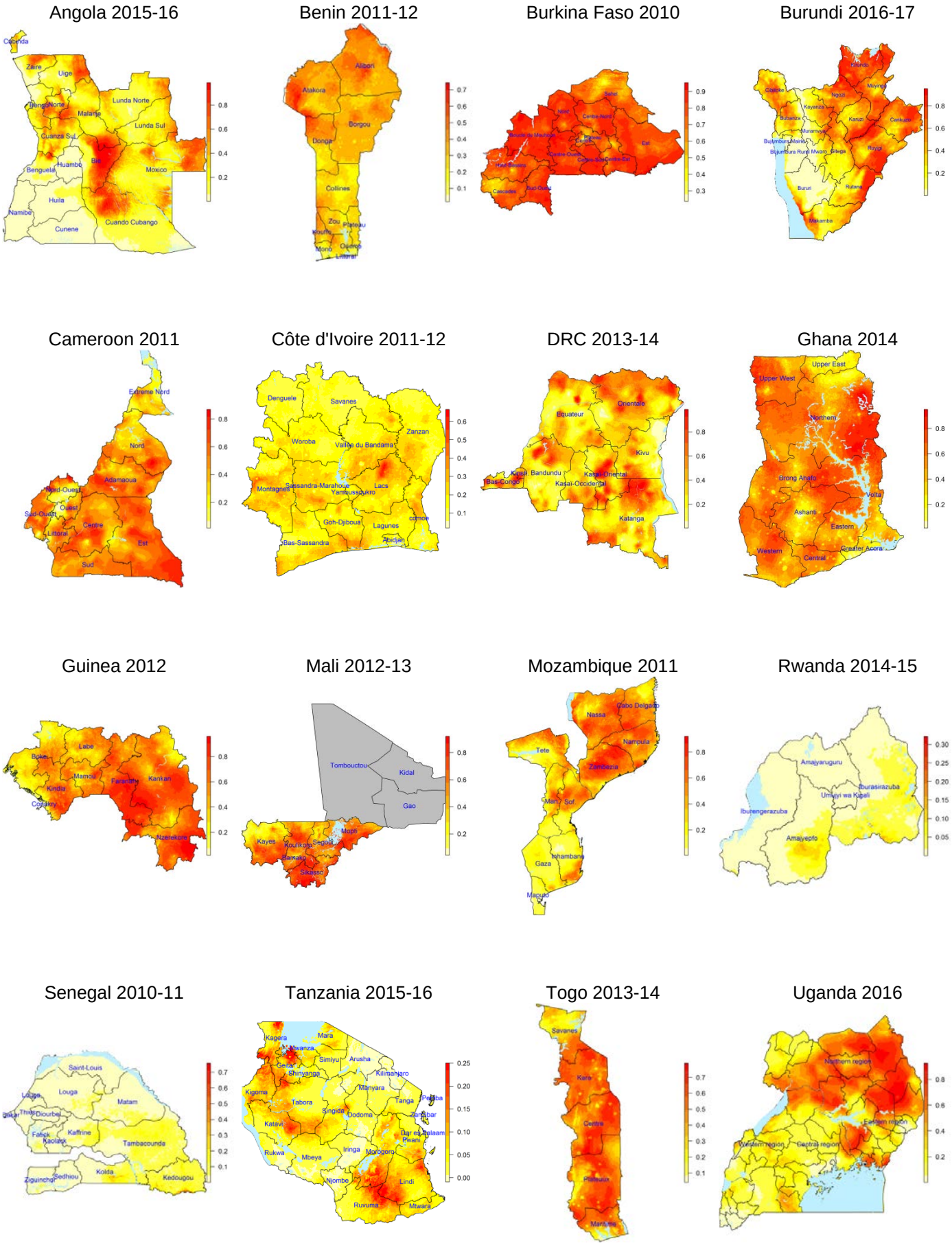


Figure 2.5: Bayesian geostatistical model based moderate/severe anemia risk estimates in 16 sub-Sahara African countries at 2x2 km^2.

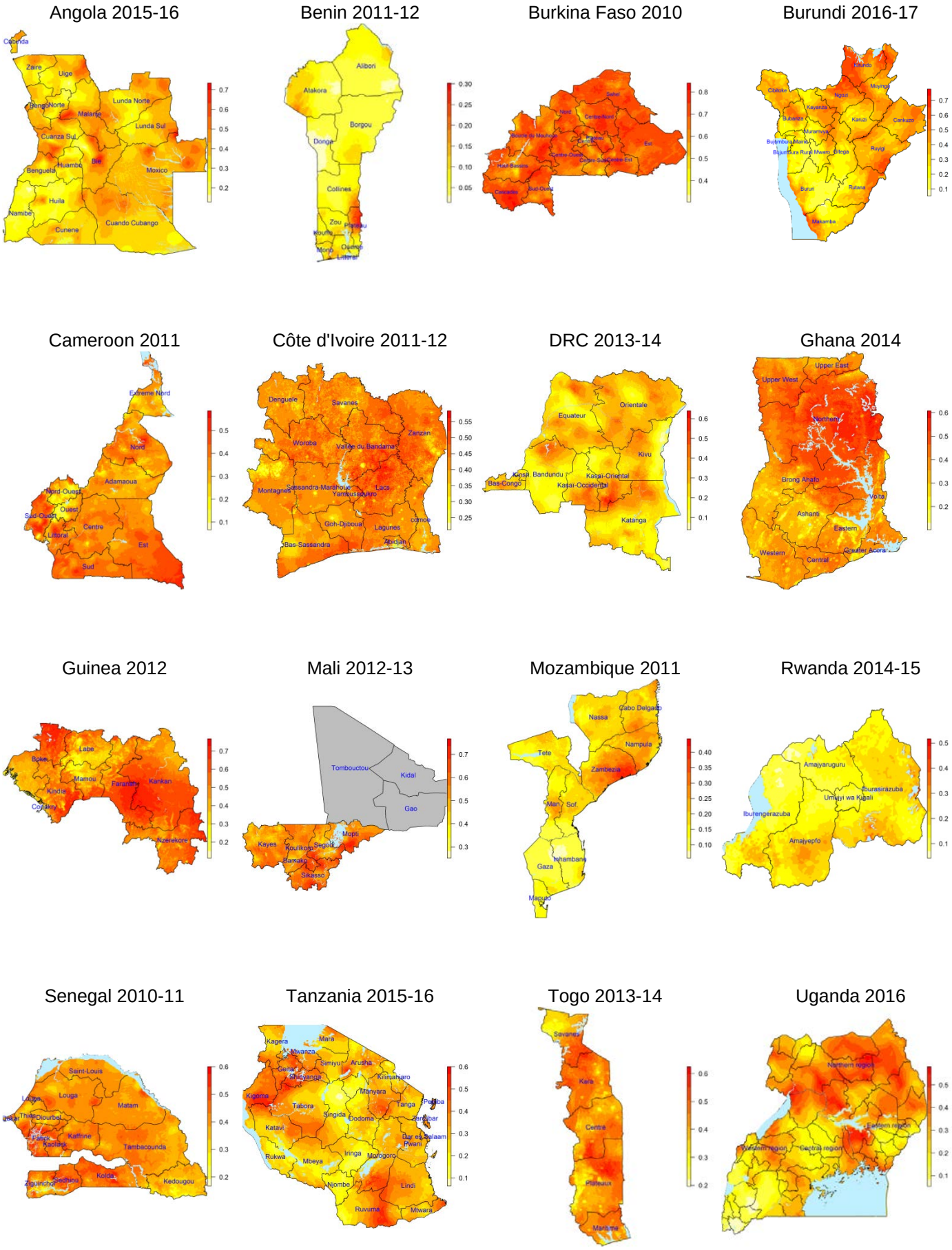


Figure 2.6: Bayesian geostatistical model based severe anemia risk estimates in 16 sub-Sahara African countries at 2x2 km^2.

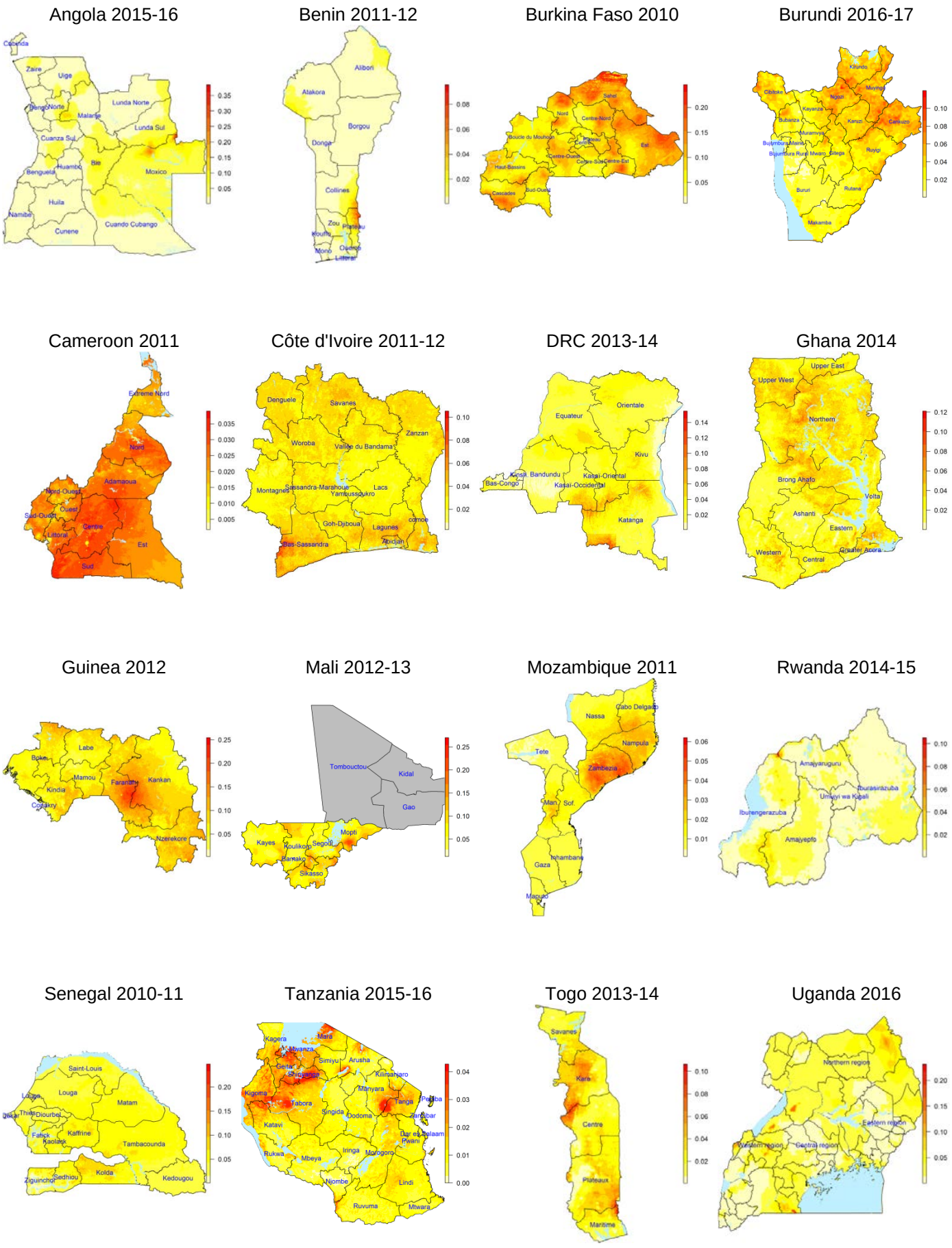
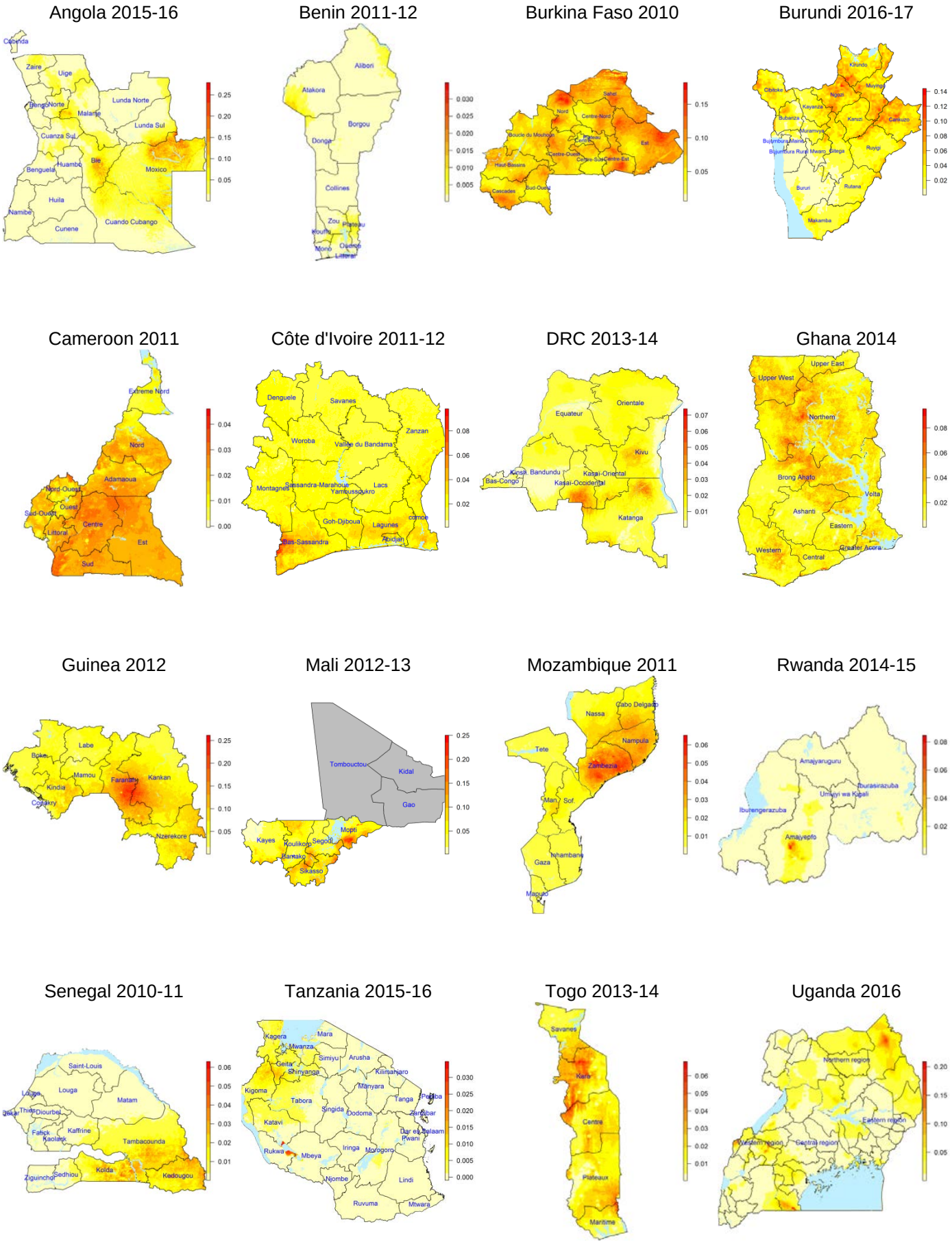


Figure 2.7: Bayesian geostatistical model based severe comorbidity risk estimates in 16 sub-Sahara African countries at 2x2 km^2.



Chapter 3

Health inequities due to variations in the association of health interventions with under-five mortality in Africa

Isidoros Papaioannou[1,2], Rainer Sauerborn[3], Jürg Utzinger[1,2], Penelope Vounatsou[1,2]

Affiliations:

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

[3]Heidelberg Institute of Global Health, University Hospital Heidelberg, Germany

Abstract

Under-five mortality accounted for 5.4 million deaths in 2017. The African continent is the most afflicted region globally, albeit saw its under-five mortality rate being more than halved between 1990 and 2017. Most of these gains were due to scale-up of various preventive or curative health interventions. Yet, the scaling-up has been unevenly distributed across various regions of the continent, resulting on strong disparities in the health intervention coverages and in the effectiveness of the interventions at sub-national level. It follows that health inequities exist among different, geographically defined, populations. We compiled data coming from the Demographic and Health surveys and developed Bayesian, Weibull-survival, spatially varying coefficient models in order to assess the sub-national associations between various health interventions and all-cause under-five mortality across 28 African countries. We then utilized these estimates to identify health inequities within countries. Uganda, Ghana, Guinea, Benin, Senegal and Togo were, in descending order, the countries with the highest number of negative associations between health interventions and all-cause under-five mortality. Vaccination, vitamin A supplementation and deworming medication for children, as well as skilled antenatal care for pregnant woman had an important association with all-cause under-five mortality in almost all countries. Benin, Côte d'Ivoire, Guinea and Mali are the most inequitable countries in Africa, followed by Cameroon, Ghana, Liberia, Senegal, Mozambique, Zambia and Uganda. Egypt, Rwanda and Burundi were the most equitable ones. We concluded on strong health inequities across various regions of Africa and on an urgent need for further assessment of the geographical effectiveness of health interventions with a specific focus on health equity within countries.

3.1 Introduction

An outstanding reduction of under-five mortality was achieved between 1990 and the end of the Millennium Development Goals (MDGs) era, as the number of global deaths was nearly halved (United Nations, 2015). In the African setting and as of 2017, the mortality of under-five years (U5M) old children was estimated to 73.7 per thousand live births, compared to a 177.5 in the year 1990 (UNICEF, WHO, World Bank, UN DESA/Population Division, 2018). Yet, an estimated 5.4 million deaths in children below the age of five occurred worldwide in 2017, with the African region being the most afflicted (WHO, 2019c). A recent report (United Nations, 2017a) from the United Nations (UN) highlighted the need for accelerated progress towards U5M reduction, if the African continent is to achieve the 3.2 target of the Sustainable Development Goals (SDGs), which champions the reduction of U5M to a maximum of 25 per thousand live birth for all countries worldwide. In Africa, the main drivers of these deaths are malaria, pneumonia, diarrhoea, birth asphyxia and preterm birth complications. Nigeria, Democratic Republic of the Congo, Ethiopia, Angola and Tanzania are among the ten countries with the highest number of pre-school deaths (Liu, 2016). Additionally, other factors such as anemia (Papaioannou, 2019) and malnutrition (WHO, 2019f), are interacting with the above causes of mortality and thus indirectly contributing to the total number of under-five deaths.

Health interventions constitute the cornerstone in reducing U5M. During the MDGs era, the coverage of health interventions has been increased across most regions in Africa, leading to unpresented mortality reduction during that time. For instance, in 2016 it was estimated that 54% of the population at risk of malaria slept under an insecticide-treated net (ITN), compared to an estimated 30% in 2010. Additionally, during the same period, household ITN ownership rose to 80% from 50% while access to an ITN within a household doubled to 61% in 2016 (WHO, 2017a). The World Health Organization (WHO) depicts the importance of health interventions by informing that access to uncomplicated, reasonable-priced health interventions in Africa could have averted approximately half of the under the age of five

deaths (WHO, 2019f). A statistical analysis of cross-sectional Demographic and Health Surveys (DHS) data from Kenya 2014 found an important association of neonatal mortality with unskilled ANC provision and absence of tetanus immunization (Arunda, 2017). A modelling study using DHS data from 149 countries found that under-five years old children living in an area with full vaccination coverage have reduced mortality risk compared to those receiving no vaccinations and argued that scaling-up of measles and maternal tetanus immunization could considerably decrease U5M in sub-Saharan Africa (SSA) (McGovern, 2015). A meta-analysis of 21 studies showed that preventive vitamin A supplementation resulted to a 15% decline of all-cause U5M, albeit the sub-group meta-analysis from two African studies showed no statistical association between vitamin A supplementation and pre-school mortality (Imdad, 2011).

The remarkable reduction in U5M and increase of health intervention's coverage have been accompanied by strong geographical variations of both metrics at sub-national level. For instance, the Global Burden of Disease (GBD) 2016 study estimated that 13 countries in SSA had below 50 under-five deaths per thousand live births, while at the same time 7 countries had above 100 under-five deaths per thousand live births (GBD 2016 Mortality Collaborators, 2017). A modelling study (Golding, 2017) estimated U5M at 5x5 km grid cells in SSA and showed substantial heterogeneity in the U5M rate between and within countries, as well as unequal declines of pre-school mortality in the period 2000-2015. In terms of health interventions, data coming from national surveys such as the Malaria Indicator Survey (MIS) can quantify such variations in coverage. For instance, within Ghana, the household ownership of *long-lasting* insecticide-treated nets (LLIN) varies from 61% in Greater Accra region to 94% in the Upper East region (Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF, 2017). In an effort to assess the implications of such variability, Bayesian geostatistical analyses of the Burkina Faso 2010 DHS and the Uganda 2011 DHS showed considerate sub-national variation of the associations between U5M and child, maternal and household health interventions (Millogo, 2019; Nambuusi, 2019). Yet, in the

African setting, modelling studies associating health interventions and pre-school mortality at sub-national level are rather scarce.

Strong sub-national spatial variation of the health interventions coverage and of the association of interventions with U5M drives health inequities among the populations that experience these differences. The WHO describes health inequity as unfair or avoidable health differences, which are present among, for instance geographically defined, population groups (Solar, 2010). Importantly, health inequity is not to be confused with health inequality, as the latter does not necessarily leads to unfairness, which is included in the definition of health inequity (Dhaliwal, 2019). A review of DHS data from 54 countries examined inequities of health interventions by comparing intervention coverages, stratified by wealth index, and reported that, in the African setting, the most inequitable countries were Chad, Nigeria, Somalia, Ethiopia, Niger and Madagascar (Barros, 2012). A qualitative study examined the equity impact of Participatory Learning and Action (PLA) health intervention in women in south Asia and Africa, through 42 focus group discussions (FGDs) and reported that accessibility, relevance and the engaging format of the PLA intervention resulted in equitable effects (Morrison, 2019). A scoping review examined the health equity in the context of Health Impact Assessment (HIA) in SSA argued the need for a more systematic stratification of community subgroups (Leuenberger, 2019). To our knowledge, there is a sparsity of studies identifying health inequities based on the effectiveness of health interventions on U5M.

Our study aims to model the geographical variation in the association between health interventions and U5M in order to identify health inequities experienced by populations within countries. To achieve that, we compiled geolocated, cross-sectional, household-based survey data coming from the DHS and MIS and developed Bayesian geostatistical models, which associate health interventions with all-cause U5M at national and sub-national level. Our approach allowed us to calculate the number of statistically important associations between interventions and mortality at regional level and hence to assess if health equity of interventions exists across the regions of a given country.

3.2 Methods

Data sources

To enable our study, we downloaded, compiled, and managed data from two different sources, i.e. the DHS program and remote sensing sources.

Mortality and health interventions data

We processed all-cause under-five mortality, health interventions, demographic and socio-economic data coming from the cross-sectional, household based DHS. These are national, regional (states and departments) and residence representative surveys collecting key information on population, health and nutrition, with an overall goal to assist countries with monitoring and evaluation queries. The DHS program uses a stratified two-stage cluster design to extract the survey samples and a variety of survey instruments to obtain the data, i.e. questionnaires, biomarkers and geographical information. Protection of the privacy of respondents and household members is achieved by the strict standards employed by the DHS program. Specifically, an informed consent statement is read to the respondent, before each interview or biomarker test, and who may accept or decline to participate. A child's parent/guardian/caregiver is responsible for providing verbal informed consent on behalf of children less than 5 years before any test is conducted. A prescribed statement to the respondent is read as a query for verbal consent and the reply for consent or assent is recorded in the questionnaire by the interviewer. The interviewer signs his or her name attesting having read the consent statement to the respondent. Written consent is not included (https://dhsprogram.com/publications/publication-dhsm7-dhs-questionnaires-and-manuals.cfm). We utilized surveys conducted between 2010 and 2017, containing information on 336,735 under-five children.

For each country and survey, we extracted a number of individual level mortality and health-related data, as well as a number of maternal, demographic and socio-economic characteristics, spanning from 5,046 individual level data for Namibia to 31,225 for Nigeria. Specifically, we had available information on the current or death age of each child, its sex,

birth order, the geographical coordinates of its residence and mode of delivery. Additionally, we had data on the educational level of each child's mother, age at birth and residence. Each child observation was linked with a wealth index, based on the wealth quantile that its household belonged to. For our work, we considered and extracted data on the following intervention categories: malaria bednets, Water, Sanitation and Hygiene (WASH) practises, reproductive health, breastfeeding, vaccinations, micronutrients supplementation and treatment of diseases. We provide a complete list and description of all health interventions in Table 4.1. We aggregated the health interventions at cluster level and linked them with each child based on the geographical coordinates of the cluster that each child was declared to reside. The aggregation method was preferred as various interventions are not reported by the DHS for the dead children and thus individual level inputs for the health interventions, in the statistical models employed by our study, would have resulted on biased estimates.

Environmental and climatic data

To adjust our models for potential confounding from infectious diseases, we used climatic and environmental data to approximate them. We compiled open-access data from the Moderate Resolution Imaging Spectroradiometer (MODIS), the U.S. Geological Survey-Earth Resources Observation Systems (USGSS) Data portal, the Shuttle Radar Topography Mission (SRTM) and the Global Rural and Urban Mapping Project (GRUMP). Specifically, we had available information on the altitude (SRTM), normalized difference vegetation index (NDVI) (MODIS), land surface temperature at night (LSTN) (MODIS), land surface temperature at day (LSTN) (MODIS), rainfall (USGSS) and the distance to the nearest water bodies (DWATER) (MODIS) for each under-five child based on its geolocations. We also had data on the land cover type (LC) (MODIS), which we recoded to correspond to one of the following categories: croplands, grasslands and forests. Based on GRUMP, we classified the residence of each child to either urban or rural. A list of environmental and climatic data used in the study, together with their spatial resolution, is provided in the Appendix.

Statistical analysis

We developed geostatistical Weibull survival models in order to associate health interventions with U5M and to identify health inequities across 28 SSA countries. Specifically, for each child there was available information on its survival status, i.e. zero (death) or one (alive), together with its death or current age. The geostatistical component of the model corresponds to a random intercept that accounts for the spatial correlation of the outcome, i.e. under-five mortality, across the different unique geolocations for which we had available data. This spatial random intercept was modelled via a cluster level Gaussian process with Matérn covariance matrix (Banerjee, 2014). The Bayesian inference summaries regression coefficients by their posterior distribution and thus implies that a coefficient is considered as statistically significant if its posterior Bayesian Credible Interval does not include the value zero. Throughout the paper, we call (statistically) important effects or associations the regression coefficients that would have been considered as statistically significant by frequentist inference.

We conducted a bivariate geostatistical variable selection procedure for each country and health intervention, separately, in order to consider interventions for the final model fit (Appendix). Then, a spatially varying coefficient extension of the Bayesian geostatistical Weibull survival model was employed and fitted for each country and intervention, separately, to estimate the effect of each intervention at sub-national (province) level. The spatially varying coefficient formulation was achieved by incorporating an additional random slope with exchangeable structure (for the coefficient of the health intervention) to the random part of the linear predictor. The posterior mean of the regression coefficient of the health intervention was interpreted as the national effect in the intervention-U5M association. A posterior sample of size 1,000 was obtained from the province-specific effects and was added up to national estimate in order to obtain the province-level association between the health intervention and U5M. When the 95% of the posterior sample was important (did not include zero), we concluded an important association. All spatially varying coefficient models

were adjusted for maternal, individual-child, socio-economic, environmental and climatic factors. We also used, prior to model fitting, the Variance Inflation Factor (VIF) to exclude highly correlated predictors from the model and thus to avoid collinearity (O'Brien, 2007). All national estimates are presented in Table 4.2 and the total number of province-level important association for each country is illustrated in Figure 4.1. In Figure 4.2, we extracted all province-level U5M rates from the DHS reports and illustrated them in a continental Africa map.

We used the R software (version 3.6.1) (R Core Team (2019), 2019) and its associated integrated development environment, the RStudio (RStudio Team (2018), 2018), for the data management, statistical analysis and mapping efforts. Bayesian inference was performed using the Integrated Nested Laplace Approximation (INLA) (Rue, 2009; Lindgren, 2011). Appendix contains further information concerning our model specification.

3.3 Results

Our study included 16,693 unique locations across 28 countries in Africa, ranging from 214 locations in Senegal to 1,741 locations in Egypt. The earliest survey incorporated in our analysis was from the Burkina Faso 2010 DHS while the latest from the Benin 2017 DHS, with the average survey year being between 2013 and 2014. Nigeria had the richest dataset, including 31,225 under-five children, while the Namibian dataset had the lowest number of observations, accounting for 5,046 children. Egypt had by far the lowest U5M rate among the countries considered in the study, with Rwanda, Senegal, Kenya and Namibia having an U5M rate of approximately 52 deaths per 1,000 live births. Contrarily, Sierra Leone, Chad, Burkina Faso, Guinea and Cameroon had, in descending order, the highest U5M. Considerable variation in coverage was observed between countries and across different health interventions. Specifically, across all countries, the highest mean coverage, accounting for more than 85%, was observed in Bacillus Calmette–Guérin (BCG) vaccination, presence of iodized salt in a household and the reception of antenatal care (ANC) from a skilled provider for pregnant woman. The lowest mean coverages, accounting

for less than 30%, were reported for iron supplementation, intake of Artemisinin-combination therapy (ACT) for children with fever and the percentage of new-borns receiving first postnatal check-up from a skilled provider within hours after delivery. When focusing on a single health intervention, strong between-countries differences in coverage were also noted. For instance, ownership of at least one Insecticide-Treated net (ITN) varied from below 35% in Angola and Namibia to above 90% in Benin. Additionally, prevalence of vitamin A supplementation for children was measured at 6% for Angola and at 86.4% for Rwanda, while exclusive breastfeeding practises ranged from 0.3% in chad to 87.3% in Rwanda. Overall, Benin championed all malaria bed-net coverages, primarily because the Benin survey was the most recent we included in our study, and Rwanda was the country with the highest coverages across most health interventions. A description of health interventions can be found in Table 4.1, while a summary of the data in the Appendix. Results from the bivariate geostatistical variable selection and the national estimates of the association between health interventions and U5M (using the spatially varying coefficient models) demonstrated substantial differences across health interventions and countries. Diphtheria, Tetanus, and Pertussis (DPT), Measles and BCG vaccinations were negatively statistically associated with all-cause U5M in all 28 countries included in our analysis. Additionally, vaccination interventions had, in most occasions, the highest magnitude of association with U5M. Vitamin A supplementation for children, skilled ANC for pregnant woman, deworming medication in the past 6 months and immunization against polio and neonatal Tetanus were associated with decreased mortality risk in more than 20 countries. Contrarily, ACT and exclusive breastfeeding had the lowest number of bivariate statistical associations. Bed-net interventions against malaria were considered in our study through five distinct coverage indicators. We observed that cumulatively in nine countries at least one bed-net indicator was associated with U5M. Uganda, Ghana, Guinea, Benin, Senegal and Togo were, in descending order, the countries with the highest number of health interventions associated with all-cause U5M. On the other hand, Tanzania had the fewest associations. On average, 12 health interventions were associated with mortality for each

country. All results from the bivariate geostatistical variable selection are presented in the Appendix.

Our multivariate geostatistical models with spatially varying coefficients for the effects of intervention coverage suggest strong sub-national health inequities, based on the number of within-country associations between health interventions and U5M. Benin, Côte d'Ivoire, Guinea and Mali are the most inequitable countries, as they contain simultaneously at least one province with very few (0 to 3) important interventions and at least one province with many effective interventions (at least 10). Cameroon, Ghana, Liberia, Senegal, Mozambique, Zambia and Uganda demonstrate also important disparities within their borders. Contrarily, Egypt, Rwanda and Burundi are the most equitable countries, as all theirs provinces share approximately the same number of health interventions with statistically important effects on U5M. Generally, poor performance of health interventions, by means of the number of statistically important associations, is observed in most parts of Angola, Chad and Nigeria, some regions of Tanzania and Ethiopia, the Savanes region of Côte d'Ivoire, the Donga department of Benin, the Mopti region in Mali and the north-west areas of Guinea (Figure 4.1).

3.4 Discussion

To our knowledge, this is the first study to link health inequity with the geographical variation of the effects of health interventions on mortality across most countries in Africa. We employed rigorous geostatistical models and used data of highest quality generated from DHS and MIS. Our results indicate health inequities based on large differences of the total number of effective interventions at subnational level, across different regions of Africa.

We defined inequalities according to the differences in the number of interventions that are statistically associated to U5M between regions. Hence, the provinces of countries like Angola are not considered to experience inequalities as they were identified having almost the same number of important associations, albeit unacceptably low. Our model-based health equity map (Figure 4.1) coincides with a recent map presenting the estimated equity

index for deworming, based on DHS data for household wealth index and the corresponding coverage, in low and middle-income countries (Lo, 2019). Another study (Wariri, 2019) utilized data from the World Health Organization (WHO) and the United Nations International Children's Emergency Fund (UNICEF) to evaluate equality gaps in Bacille Calmette Guerin (BCG) and diphtheria-tetanus-pertussis dose 3 (DPT3) and reported health inequities for Guinea, Mali, Nigeria (BCG and DPT3) and Cote d'Ivoire (BCG), which we also reported among the most inequitable countries. A recent trend and multi-country pooled analysis (Taylor, 2017) studied equity tendencies in ITN ownership using DHS data for 19 SSA countries and concluded that Mali and Mozambique had not experience better equity from 2006-2007 to 2011-2013 (years in which the DHS surveys were conducted), with Senegal showing worsened equity trends among the two periods. These results from the trend analysis coincide with our conclusions, albeit they also reported a decrease in equity for Angola.

Our statistical analysis models associations between the outcome and the predictors and it does not imply causal relations, for which a causal modelling approach should have been used. Lack of association between health interventions and U5M, as obtained from our variable selection approach, does not imply that the corresponding intervention is not effective against U5M, but instead it means that geographical changes in the coverage of a specific intervention are not associated with changes in all-cause under-five mortality. Our analysis does not consider changes in the coverage of health interventions during the 5 years period that mortality has been estimated from a given survey. This may have led to an underestimation of the association of some of the interventions with all-cause pre-school mortality.

Acknowledgements

The authors are grateful to the DHS for making the data available and all collaborating institutions whose contributions have enabled this study. The Swiss National Science

Foundation (SNSF) Swiss Programme for Research on Global Issues for Development (R4D) project no. IZ01Z0-147286 funded this work.

Ethical approval and consent to participate

In this study we analysed secondary data made available by the Demographic Health Survey (DHS) MEASURE. According to survey protocols and related documents of the surveys, ethical approval was obtained from the Institutional Review Board of International Consulting Firm (ICF) of Calverton, Maryland, USA, and from the national ethical committees in the countries that the surveys were contacted. Details of ethical clearance are published in the DHS reports available at https://dhsprogram.com/publications/index.cfm.

Competing interests

We declare no competing interests.

Role of funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the paper. The authors had full access to all data in the study and had final responsibility for the decision to submit for publication.

Contributors

IP conceptualised the project, processed and analysed the data, interpreted the results, and wrote the manuscript. PV acquired the financial support of the project, conceptualised the project and assisted in statistical analysis. RS conceptualised the project. RS, JU and PV revised the manuscript and provided important intellectual content. All authors approved the final version of the manuscript.

Availability of Materials and Data

The study data are available upon request from the Demographic and Health Surveys program (https://dhsprogram.com/data/Access-Instructions.cfm). The “Dataset Terms of Use” does not allow us to pass on downloaded data.

Table 3.1: Description of health interventions.

Intervention	Description of Intervention
Malaria bednets	
%H_1ITN	Percentage of households with at least one Insecticide-Treated net (ITN)
%H_1ITN2	Percentage of households with at least one ITN for every two people
%P_ITNA	Percentage of population with access to an ITN within their household
%P_ITNS	Percentage of population in a household that slept under an ITN the previous night before the survey
%P_ITN5	Percentage of under-five children in a household who slept under an ITN the night before the survey
%P_ITNU	Percentage of existing ITNs used by the population in a household the previous night of the survey
WASH	
Improved Water	Percentage of households with improved source of drinking water
Improved Sanitation	Percentage of households with improved sanitation facilities
Reproductive health	
ANC provider	Percentage of pregnant mothers receiving antenatal care (ANC) from a skilled provider
4+ ANC visits	Percentage of pregnant woman making at least four ANC visits during their pregnancy
Fansidar	Percentage of pregnant woman which received Fansidar during pregnancy
Postnatal Care	Percentage of new-borns receiving first postnatal check-up from a skilled provider within hours after delivery
Breastfeeding	
Immediate	Percentage of infants which breastfed within one hour after birth
Exclusive	Percentage of infants which exclusively breastfed during the first six months after delivery
Vaccinations	
Tetanus	Percentage of last-born children protected against neonatal Tetanus
BCG	Percentage of children vaccinated against Bacillus Calmette–Guérin (BCG)
DPT	Percentage of children with complete Diphtheria, Tetanus, and Pertussis (DPT) vaccination
Polio	Percentage of children with complete Polio vaccination
Measles	Percentage of children vaccinated against Measles
Micronutrients	
Vitamin A	Percentage of children which received vitamin A supplements in the past 6 months
Iron	Percentage of children which received iron supplements in the past 7 days
Iodized salt	Percentage of children which lived in a household with positive test for iodized salt
Treatments	
ORS	Percentage of children with diarrhoea given fluid from oral rehydration solution (ORS)
ACT	Percentage of children with fever during the two weeks prior to the survey which took Artemisinin-combination therapy (ACT)
Deworming	Percentage of children which took deworming medication in the past 6 months

Table 3.2: Posterior estimates of health interventions on U5M, at national scale, obtained from a spatially varying coefficient model and adjusted for demographic, socio-economic and climatic confounders. Estimates are presented for the health interventions, which were estimated to be statistically associated with U5M, based on a bivariate geostatistical variable selection procedure.

	Angola	Benin	Burkina Faso	Burundi	Cameroon	Chad	DRC
H_1ITN							
H_1ITN2							
P_ITNA							
P_ITNS							
P_ITN5							-0.43 (-0.85,-0.01)
P_ITNU							-0.68 (-1.23,-0.14)
Improved Water			-0.33 (-0.60,-0.06)				
Improved Sanitation		-0.59 (-1.29,0.10)	-0.33 (-0.90,0.24)	-0.59 (-0.99,-0.19)		-0.88 (-1.47,-0.28)	-0.54 (-0.92,-0.16)
ANC provider	-0.63 (-1.30,0.03)	-0.80 (-1.33,-0.26)	-0.46 (-0.86,-0.07)	-2.27 (-3.38,-1.15)	-0.61 (-1.25,0.02)	-0.52 (-0.94,-0.10)	
4+ ANC visits	-0.28 (-0.96,0.40)	-0.81 (-1.38,-0.25)	-0.23 (-0.79,0.21)		-1.01 (-1.58,-0.43)		
Fansidar	-0.27 (-0.96,0.43)	-1.03 (-1.54,-0.51)				-0.59 (-1.12,-0.05)	
Post. Care			-0.44 (-1.01,0.13)				
Immediate					-0.25 (-0.65,0.16)		
Exclusive							
Tetanus	-0.68 (-1.36,0)	-1.07 (-1.59,-0.55)	-1.54 (-2.16,-0.92)		-0.86 (-1.50,-0.23)	-0.50 (-0.92,-0.07)	
BCG	-0.73 (-1.37,-0.08)	-1.82 (-2.45,-1.19)	-3.12 (-3.61,-2.62)	-4.02 (-4.92,-3.13)	-3.21 (-3.77,-2.64)	-0.65 (-0.94,-0.36)	-1.43 (-1.76,-1.10)
DPT	-1.16 (-2.02,-0.30)	-1.54 (-2.13,-0.95)	-1.82 (-2.24,-1.41)	-3.54 (-4.42,-2.67)	-1.62 (-2.04,-1.21)	-0.52 (-0.90,-0.14)	-1.09 (-1.42,-0.76)
Polio	-0.77 (-1.68,0.14)	-1.37 (-1.98,-0.76)	-1.56 (-1.97,-1.16)	-2.47 (-3.23,-1.71)	-1.18 (-1.60,-0.75)	-0.59 (-1.06,-0.13)	-0.80 (-1.27,-0.34)
Measles	-0.69 (-1.50,0.12)	-1.37 (-2.03,-0.71)	-2.16 (-2.60,-1.71)	-2.92 (-3.77,-2.08)	-2.06 (-2.53,-1.58)	-0.69 (-1.01,-0.37)	-1.51 (-1.90,-1.13)
Vitamin A		-0.75 (-1.13,-0.37)	-0.84 (-1.18,-0.51)	-1.50 (-1.99,-1.02)	-0.83 (-1.28,-0.37)	-0.48 (-0.86,-0.10)	-0.94 (-1.28,-0.59)
Iron		-0.40 (-0.92,0.11)			-1.03 (-2.06,-0.01)		-1.06 (-1.62,-0.50)
Iodized Salt		-0.82 (-1.47,-0.17)					
ORS					-1.76 (-3.49,-0.03)		
ACT						-0.88 (-1.75,-0.01)	
Deworming		-0.79 (-1.29,-0.29)	-0.43 (-1.00,0.13)	-1.53 (-2.07,-0.99)	-1.57 (-2.09,-1.04)	-0.19 (-0.61,0.22)	-0.92 (-1.32,-0.53)

	Egypt	Ethiopia	Gabon	Ghana	Guinea	Ivory Coast	Kenya
H_1ITN					-0.62 (-1.18,-0.06)		
H_1ITN2							
P_ITNA					-0.70 (-1.52,0.11)		
P_ITNS							
P_ITN5				-0.93 (-1.77,-0.10)			
P_ITNU							
Improved Water							
Improved Sanitation	-0.34 (-0.87,0.20)	-0.55 (-1.43,0.33)			-0.50 (-1.11,0.12)	-0.29 (-0.85,0.27)	
ANC provider	-0.46 (-1.19,0.27)	-0.41 (-0.94,0.11)	-1.98 (-3.03,-0.92)	-2.80 (-4.05,-1.56)	-0.85 (-1.40,-0.30)	-1.58 (-2.26,-0.90)	-1.36 (-1.98,-0.74)
4+ ANC visits		-0.56 (-1.22,0.09)		-1.51 (-2.44,-0.57)	-0.80 (-1.45,-0.15)	-0.83 (-1.50,-0.16)	
Fansidar				-1.93 (-2.92,-0.94)	-0.26 (-0.91,0.39)		
Post. Care				-1.29 (-2.43,-0.15)			
Immediate		-0.47 (-0.96,0.02)	-0.87 (-1.52,-0.22)	-0.76 (-1.41,-0.11)		-0.34 (-0.82,0.14)	-0.67 (-0.27,0.03)
Exclusive							
Tetanus	-1.12 (-1.70,-0.55)	-0.74 (-1.29,-0.19)	-1.95 (-3.04,-0.85)	-1.91 (-3.02,-0.81)	-1.10 (-1.74,-0.45)	-1.37 (-2.03,-0.71)	
BCG	-6.14 (-6.57,-5.71)	-0.99 (-1.61,-0.36)	-2.84 (-3.70,-1.97)	-4.78 (-5.84,-3.73)	-1.65 (-2.18,-1.13)	-1.85 (-2.34,-1.37)	-5.58 (-6.22,-4.94)
DPT	-4.61 (-5.13,-4.10)	-0.61 (-1.29,0.07)	-1.62 (-2.60,-0.64)	-2.38 (-3.24,-1.52)	-0.62 (-1.12,-0.13)	-1.28 (-1.75,-0.81)	-2.90 (-3.38,-2.42)
Polio	-3.86 (-4.41,-3.31)	-1.20 (-2.17,-0.23)		-1.49 (-2.25,-0.74)	-0.84 (-1.40,-0.28)	-1.06 (-1.56,-0.57)	-0.88 (-1.25,-0.51)
Measles	-1.90 (-2.47,-1.33)	-1.05 (-1.77,-0.32)	-1.55 (-2.42,-0.68)	-3.53 (-4.55,-2.50)	-1.11 (-1.62,-0.60)	-1.40 (-1.96,-0.84)	-2.74 (-3.27,-2.22)
Vitamin A	-1.25 (-2.05,-0.46)	-1.01 (-1.50,-0.52)	-1.75 (-2.61,-0.89)	-1.26 (-2.01,-0.50)	-0.72 (-1.27,-0.16)	-0.47 (-0.92,-0.02)	-1.59 (-2.01,-1.17)
Iron	-1.39 (-2.59,-0.20)						
Iodized Salt							
ORS					-1.88 (-3.58,-0.18)		
ACT				-0.61 (-1.05,-0.16)			
Deworming	-1.36 (-2.43,-0.28)	-1.50 (-2.43,-0.57)	-1.71 (-2.57,-0.85)	-1.83 (-2.82,-0.84)	-1.55 (-2.33,-0.76)	-0.88 (-1.44,-0.32)	-1.66 (-2.32,-1.00)

	Liberia	Malawi	Mali	Mozambique	Namibia	Nigeria	Rwanda
H_1ITN							-1.00 (-1.93,-0.07)
H_1ITN2							
P_ITNA							-0.81 (-1.68,0.06)
P_ITNS		-0.58 (-1.11,-0.04)	-1.00 (-1.76,-0.24)				
P_ITN5		-0.76 (-1.20,-0.33)	-0.82 (-1.44,-0.20)				
P_ITNU		-0.44 (-0.92,0.03)					
Improved Water						-0.12 (-0.28,0.04)	
Improved Sanitation						-0.17 (-0.40,0.05)	
ANC provider	-1.71 (-2.66,-0.77)	-1.91 (-2.67,-1.14)	-0.96 (-1.60,-0.32)	-1.04 (-1.76,-0.32)	-1.31 (-2.46,-0.16)	-0.19 (-0.53,0.14)	-1.99 (-3.36,-0.63)
4+ ANC visits	-1.20 (-1.91,-0.49)	-0.83 (-1.43,-0.23)	-1.22 (-1.96,-0.49)	-0.91 (-1.54,-0.27)	-0.84 (-1.74,0.06)	-0.08 (-0.44,0.28)	-1.57 (-2.48,-0.65)
Fansidar		-1.87 (-2.59,-1.15)		-1.02 (-1.70,-0.35)		-0.19 (-0.67,0.29)	
Post. Care						-0.74 (-1.46,-0.03)	
Immediate		-0.85 (-1.35,-0.34)		-0.63 (-1.15,-0.11)	-1.16 (-1.95,-0.36)	-0.46 (-0.75,-0.17)	-1.66 (-2.58,-0.73)
Exclusive							
Tetanus	-1.63 (-2.44,-0.83)	-1.42 (-2.09,-0.75)	-0.91 (-1.54,-0.27)	-0.92 (-1.55,-0.30)		-0.24 (-0.60,0.11)	
BCG	-2.31 (-2.89,-1.73)	-2.94 (-3.70,-2.19)	-1.50 (-1.96,-1.04)	-3.25 (-3.81,-2.67)	-5.62 (-6.79,-4.45)	-0.93 (-1.22,-0.63)	-8.12 (-9.47,-6.78)
DPT	-1.44 (-1.97,-0.90)	-2.24 (-2.98,-1.51)	-0.90 (-1.36,-0.45)	-1.96 (-2.44,-1.47)	-2.36 (-3.21,-1.51)	-1.00 (-1.33,-0.68)	-5.38 (-6.41,-4.35)
Polio	-1.03 (-1.63,-0.43)	-0.91 (-1.53,-0.28)	-0.69 (-1.28,-0.10)	-1.59 (-2.05,-1.12)	-1.39 (-2.22,-0.56)	-0.94 (-1.32,-0.55)	-2.81 (-3.55,-2.07)
Measles	-2.06 (-2.67,-1.45)	-1.37 (-2.12,-0.63)	-1.12 (-1.61,-0.64)	-2.72 (-3.25,-2.18)	-2.64 (-3.56,-1.72)	-0.97 (-1.31,-0.64)	-0.77 (-1.38,-0.16)
Vitamin A	-0.87 (-1.36,-0.37)	-1.50 (-1.98,-1.03)	-1.01 (-1.45,-0.57)	-2.27 (-2.80,-1.75)	-2.83 (-3.64,-2.02)	-0.91 (-1.24,-0.58)	-3.31 (-4.18,-2.44)
Iron	-0.68 (-1.47,0.11)		-1.24 (-1.84,-0.64)	-0.67 (-1.36,0.01)			
Iodized Salt		-0.33 (-0.64,-0.01)					
ORS				-1.49 (-2.87,-0.11)			
ACT							
Deworming	-0.73 (-1.32,-0.14)	-1.50 (-2.01,-1.00)	-0.70 (-1.22,-0.19)	-1.36 (-1.90,-0.82)	-1.13 (-1.93,-0.35)	-0.68 (-1.10,-0.26)	-3.24 (-4.11,-2.37)

	Senegal	Sierra Leone	Tanzania	Togo	Uganda	Zambia	Zimbabwe
H_1ITN					-1.02 (-1.54,-0.50)		
H_1ITN2					-0.75 (-1.25,-0.25)		
P_ITNA					-0.95 (-1.46,-0.44)		
P_ITNS					-0.72 (-1.23,-0.22)	-0.75 (-1.31,-0.19)	
P_ITN5					-0.66 (-1.10,-0.22)	-0.59 (-1.04,-0.15)	
P_ITNU	-0.71 (-1.29,-0.13)						
Improved Water				-0.28 (-0.64,0.07)			
Improved Sanitation	-0.33 (-1.03,0.36)			-1.08 (-2.44,0.28)			
ANC provider	-0.79 (-2.07,0.48)	-1.18 (-2.19,-0.18)		-0.70 (-1.39,-0.02)	-1.55 (-2.43,-0.68)	-1.54 (-2.38,-0.72)	-2.39 (-3.22,-1.55)
4+ ANC visits	-1.73 (-2.84,-0.62)	-0.83 (-1.40,-0.26)		-0.62 (-1.36,0.13)	-0.88 (-1.53,-0.24)		-1.76 (-2.60,-0.91)
Fansidar	-1.22 (-2.39,-0.05)			-1.18 (-1.92,-0.44)	-1.06 (-1.73,-0.39)	-1.29 (-2.01,-0.56)	
Post. Care	-1.39 (-2.28,-0.50)						
Immediate				-0.54 (-1.03,-0.05)		-0.57 (-1.02,-0.12)	-0.61 (-1.37,0.15)
Exclusive						-1.62 (-2.90,-0.35)	-1.19 (-2.11,-0.26)
Tetanus	-0.79 (-1.96,-0.38)	-1.42 (-2.37,-0.48)		-0.91 (-1.75,-0.08)	-1.18 (-1.90,-0.46)		-2.21 (-3.08,-1.35)
BCG	-1.86 (-2.80,-0.93)	-3.63 (-4.27,-3.00)	-2.04 (-2.89,-1.19)	-3.71 (-4.57,-2.85)	-2.81 (-3.60,-2.01)	-4.57 (-5.21,-3.92)	-1.79 (-2.79,-0.79)
DPT	-1.74 (-2.51,-0.96)	-1.57 (-2.03,-1.11)	-1.28 (-2.05,-0.52)	-1.76 (-2.40,-1.13)	-2.23 (-2.91,-1.56)	-2.34 (-2.86,-1.82)	-1.07 (-2.06,-0.08)
Polio	-1.01 (-1.76,-0.26)	-1.54 (-1.99,-1.08)		-1.37 (-2.08,-0.67)	-1.90 (-2.57,-1.22)	-1.04 (-1.59,-0.49)	
Measles	-2.03 (-2.86,-1.20)	-2.29 (-2.88,-1.70)	-1.28 (-2.14,-0.43)	-1.96 (-2.67,-1.26)	-1.86 (-2.60,-1.11)	-3.34 (-3.95,-2.74)	-1.81 (-2.88,-0.73)
Vitamin A	-1.78 (-2.74,-0.83)	-1.92 (-2.43,-1.42)	-1.03 (-1.64,-0.42)	-2.30 (-2.91,-1.69)	-1.08 (-1.54,-0.62)	-2.10 (-2.62,-1.59)	-2.10 (-2.62,-1.59)
Iron		-1.00 (-1.52,-0.49)					
Iodized Salt		-0.36 (-0.70,-0.01)			-0.65 (-1.54,0.24)		
ORS			-2.63 (-4.32,-0.94)				
ACT							
Deworming	-1.61 (-2.54,-0.67)	-1.03 (-1.58,-0.48)	-1.02 (-1.64,-0.40)	-0.74 (-1.32,-0.15)	-1.40 (-1.89,-0.92)	-1.62 (-2.15,-1.10)	

Figure 3.1: Number of statistically important associations between health interventions and U5M, presented at administrative level 1, based on a spatially varying coefficient modelling approach.

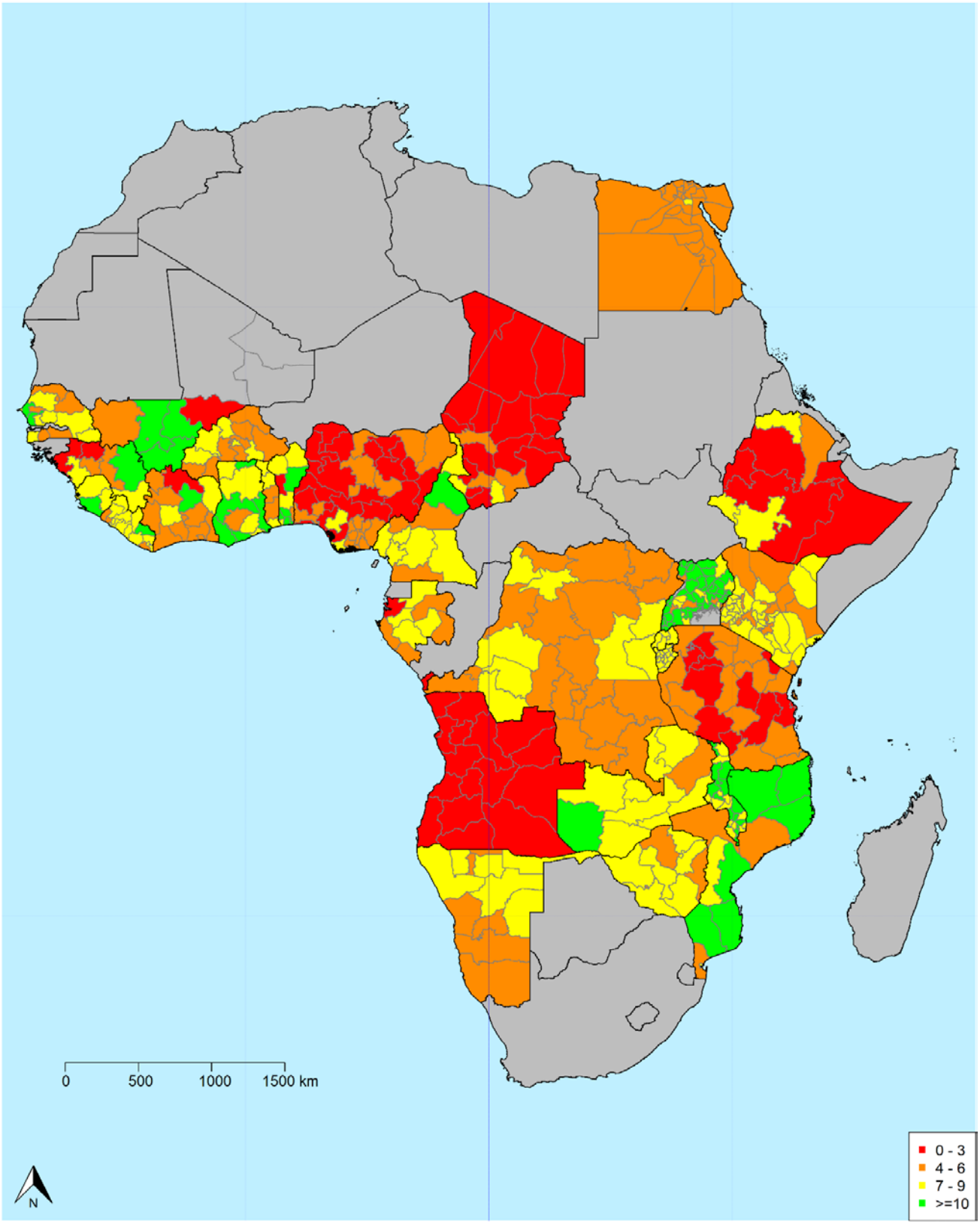
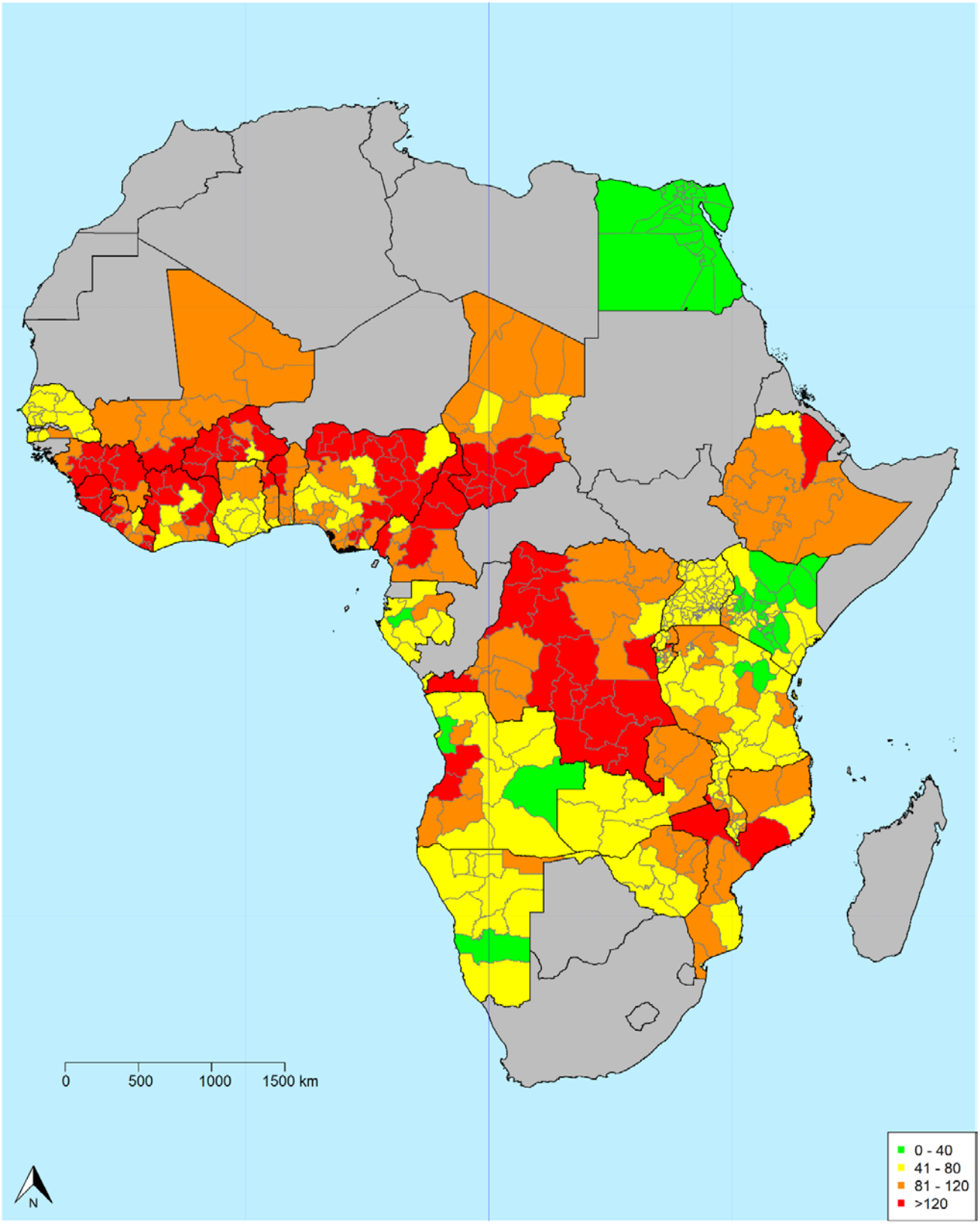


Figure 3.2: Under-five mortality rate, presented at administrative level 1, for all countries included in the study. Data were extracted from the DHS program.



Egypt: All areas have an U5M rate of less than 40 per thousand live births, except the rural regions of Upper Egypt, which have an U5M rate of 42 deaths per thousand live births.
Senegal: U5M rates for the Senegal continuous DHS 2016 were reported by major regions, instead of provinces. We used data from the continuous DHS 2017, which reported U5M at province-level.
Uganda: U5M rates were reported at aggregated regions, instead of provinces, and thus we used the national average for U5M.

3.5 Appendix

Model specification

Let the age of a child being represented by T_{ji}, with j corresponding to each child and i corresponding to one of the n surveyed locations, i.e. $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$ is the set of surveyed locations. T_{ji} was assigned a Weibull lifetime distribution which assumes a fixed shape parameter α and a scale parameter λ that is reparametrized in terms of model parameters. Specifically:

$$\lambda_{ji} = \exp(\eta_{ji}),$$

$$\eta_{ji} = \beta_0 + \sum_{\kappa=1}^{K} \beta_\kappa x_{ji\kappa} + \xi_i,$$

where η_{ji} is the linear predictor (linked with the scale parameter using the log link function), $\beta = (\beta_0, \beta_1, \dots, \beta_\kappa)^T$ is the vector of K regression coefficients (corresponding to the health intervention and its confounders) and $x_{ji\kappa}$ represents the value of the κ-th predictor for the j-th child at the i-th location. A zero-mean, multivariate Gaussian distributed random vector with Matérn covariance matrix $\xi(s) = (\xi_1, \xi_2, \dots, \xi_n)^\tau$ corresponds to a location-specific (cluster) random effect that accounts for the spatial correlation in the outcome, i.e. under-five mortality. As the name suggests, the Matérn function governs the covariance between locations s_1 and s_2, specifically $\xi(s) \sim N(0, \Sigma_1)$ and $\Sigma_1(s_1, s_2) = \frac{\sigma_1^2(\kappa_1\, d(s_1,s_2))^\nu K_\nu(\kappa_1\, d(s_1,s_2))}{\Gamma(\nu)\, 2^{\nu-1}}$, with σ_1^2 being the spatial process variance, $d(s_1, s_2)$ the distance between locations s_1 and s_2 and κ_1 the scaling parameter. The modified Bessel function of second kind and order ν is represented by K_ν in the Matérn's function formula. The Matérn specification defines the spatial range r_1 as $r_1 = \frac{\sqrt{8}}{\kappa_1}$ that particularly determines the distance at which the spatial correlation becomes negligible. The above model specification describes the bivariate geostatistical Weibull survival model, which we used to select the subset of statistically important health interventions from all 25 health interventions considered in our study. The selected interventions were fitted separately for each country using an extension of the

above model, specifically a spatially varying coefficient model. Assuming that from the set of regression coefficients $\beta = (\beta_0, \beta_1, \dots, \beta_\kappa)^T$, the β_1 corresponds to the effect of a health interventions on under-five mortality and hence the set $(\beta_2, , \dots, \beta_\kappa)^T$ corresponds to the confounders. The spatially varying specification is achieved by the addition of a random slope, which makes the model flexible to allow for estimation of sub-national (province) level effects:

$$\lambda_{ji} = \exp(\eta_{ji}),$$

$$\eta_{ji} = \beta_0 + \left(\beta_1 + \varepsilon_{1\,m(i)}\right) x_{ji1} + \sum_{\kappa=2}^{K} \beta_\kappa x_{ji\kappa} + \xi_i,$$

where now, we β_1 represents the national effect of the health intervention on under-five mortality, $\varepsilon_1 = (\varepsilon_{11}, \dots, \varepsilon_{1M})^T$ are the spatially varying effects of β_1 at the different M provinces of each country and $m(i)$ acting as a filter on which of the M provinces the j-th child at the i-th location belongs. ε_1 is a vector of independent and Gaussian distributed random variables for which their exchangeable structure implies that region identifiers are weighted by the values of the health intervention coverage and follow an exchangeable format. The parameterization of ε_1 is as follows:

$$\pi(\varepsilon_1|\tau) = \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi}} \sqrt{\tau} \exp\left(-\frac{1}{2}\, \tau\, \varepsilon_{1m}^2\right).$$

Environmental and climatic data

Table 3.3: Sources of environmental and climatic data.

Data	Source	Spatial resolution
Annual average Normalised Difference Vegetation Index (NDVI)	MODIS	1x1 km^2
Annual average Day and Night Land Surface Temperature (LST)	MODIS	1x1 km^2
Land Cover Type (LC)	MODIS	0·5x0·5 km^2
Distance from water bodies (DWATER)	MODIS	0·5x0·5 km^2
Annual average Rainfall	USGSS	8x8 km^2
Altitude (Digital Elevation model)	SRTM	0·5x0·5 km^2
Urban rural extent	GRUMP	1x1 km^2

Additional Results

Table 3.4: Summary of the DHS data extracted from the DHS STATcompiler (part I).

	Survey year	Locations number	Children number	U5M[1]	H_1ITN[2]	H_1ITN2[2]	P_ITNA[2]	P_ITNS[2]	P_ITN5[2]	P_ITNU[2]	Improved Water	Improved Sanit.[3]	ANC provider	4+ ANC visits
Angola	2015	625	14322	68	30.9	11.3	19.7	17.6	21.7	71.0	52.3	32.2	81.6	61.4
Benin	2017	540	13262	96	91.5	60.5	77.2	71.1	76.3	73.4	71.2	13.3	83.2	52.1
Burkina Faso	2010	541	14159	129	56.9	18.5	36.1	31.5	47.4	79.6	77.0	15.2	94.9	33.7
Burundi	2016	552	13135	78	46.2	17.1	32.3	34.7	39.9	86.9	82.9	39.4	99.2	49.3
Cameroon	2011	577	11723	122	36.4	9.0	20.9	14.8	21.0	61.9	70.8	35.9	84.7	62.2
Chad	2014	624	18623	133	77.3	42.4	61.2	33.3	36.4	48.6	55.3	6.6	63.7	31.0
DRC	2013	492	17152	104	70.0	25.4	46.5	50.2	55.8	82.8	48.7	18.4	88.4	48.0
Egypt	2014	1741	15744	27	*	*	*	*	*	*	97.8	76.6	90.3	82.8
Ethiopia	2016	622	10221	88	*	*	*	*	*	*	64.8	6.3	62.4	31.8
Gabon	2012	330	6019	65	36.1	14.5	26.9	26.7	38.8	86.6	93.0	33.7	94.7	77.6
Ghana	2014	423	5801	60	68.3	45.2	59.0	35.7	46.6	48.6	89.8	13.6	97.3	87.3
Guinea	2012	300	7039	123	47.4	9.7	25.3	18.9	26.0	68.7	74.9	19.0	85.2	56.6
Ivory Coast	2011	641	7555	108	67.3	31.7	49.0	33.2	37.2	58.8	78.3	18.0	90.6	44.2
Kenya	2014	1584	20851	52	58.9	34.5	48.2	42.6	54.3	77.3	71.3	22.7	95.5	57.6
Liberia	2013	322	7606	94	54.6	22.1	37.0	31.7	38.1	67.4	72.6	14.2	95.9	78.1
Malawi	2015	850	17286	64	56.9	23.5	38.8	33.9	42.7	73.3	87.2	51.8	94.8	50.6
Mali	2012	413	10326	95	84.4	41.8	65.1	60.5	69.0	88.7	66.0	22.0	74.2	41.2
Mozambique	2011	609	11089	97	51.4	22.6	37.0	29.5	35.7	68.3	51.0	16.7	90.6	50.6
Namibia	2013	537	5046	54	24.4	12.0	18.1	3.9	5.6	21.1	86.9	33.5	96.6	62.5
Nigeria	2013	889	31225	128	49.5	22.1	36.1	12.9	16.6	34.7	60.6	30.1	60.6	51.1
Rwanda	2014	492	7856	50	80.6	42.6	63.8	61.4	67.7	77.4	72.9	54.1	99.0	43.9
Senegal	2016	214	6725	51	82.4	56.4	75.7	63.1	68.2	66.6	81.8	44.7	95.9	53.8
Sierra Leone	2013	435	11938	156	64.4	14.9	37.7	41.8	49.0	91.9	60.6	9.6	97.1	76.0
Tanzania	2015	208	10233	67	65.6	38.8	55.9	49.0	54.4	69.4	61.4	19.1	98.0	50.7
Togo	2013	329	6979	88	65.4	32.9	48.8	33.6	42.8	61.2	64.1	12.2	50.0	57.2
Uganda	2016	685	15276	64	78.4	51.1	64.6	55.0	62.0	74.0	78.3	18.7	97.3	59.9
Zambia	2013	719	13412	75	67.7	27.4	46.6	34.9	40.6	63.6	64.5	25.4	95.7	55.5
Zimbabwe	2015	399	6132	69	47.9	26.4	37.2	8.5	9.0	18.8	78.1	37.0	93.3	75.7
Mean	2013-14	625	14322	85.9	60.0	29.0	44.8	35.8	42.4	66.1	71.9	26.4	87.5	56.5

1. Under-five mortality; 2. Represents percentages; 3. Improved Sanitation
* No data

Table 3.5: Summary of the DHS data extracted from the DHS STATcompiler (part II).

	Fans.[1]	Post. Care[2]	Immed.[3]	Exclus.[4]	Tetanus	BCG	Measles	Vitamin A	Iron	Iodized Salt	ORS	ACT	Dew.[5]
Angola	55.9	16.3	48.3	37.5	66.2	71.9	56.1	6.0	*	88.5	42.6	76.7	*
Benin	50.3	60.4	54.1	41.5	70.7	88.2	67.9	52.2	23.9	91.4	22.2	6.5	38.3
Burkina Faso	73.8	*	42.1	24.8	85.5	96.5	87.3	63.0	7.4	95.4	21.2	8.7	12.8
Burundi	29.6	4.2	85.0	82.8	84.3	97.7	93.8	68.7	7.9	99.0	35.6	5.3	67.6
Cameroon	44.0	*	39.8	20.2	73.3	87.1	70.6	55.3	9.2	90.4	17.2	6.1	51.1
Chad	24.2	3.4	23.0	0.3	55.7	59.6	56.9	44.1	15.7	82.8	20.4	2.7	26.6
DRC	34.8	5.9	51.9	47.6	65.5	83.4	71.6	70.4	15.6	92.1	39.1	5.0	60.6
Egypt	*	14.2	27.1	39.7	74.4	99.1	95.8	16.7	7.6	90.1	28.4	*	10.8
Ethiopia	*	11.7	73.3	57.5	49.0	69.2	54.3	44.7	9.2	87.9	29.5	0.7	12.7
Gabon	19.5	12.1	32.3	6.0	78.3	91.5	74.3	53.8	20.5	97.4	26.1	8.8	72.3
Ghana	82.9	18.6	55.6	52.3	78.0	96.8	89.3	65.2	24.4	61.9	48.6	37.9	38.3
Guinea	31.7	16.3	16.6	20.5	75.6	82.4	61.8	40.8	11.5	65.8	34.3	4.8	28.5
Ivory Coast	29.0	26.7	30.8	12.1	67.2	83.4	64.5	60.8	13.4	90.9	17.2	3.0	36.7
Kenya	30.4	33.1	62.2	61.4	75.6	96.7	87.1	71.7	2.7	99.5	53.8	85.8	22.2
Liberia	68.1	28.8	61.2	55.2	87.8	93.9	74.2	60.2	26.7	98.7	60.4	23.9	56.1
Malawi	90.5	58.1	76.3	61.0	90.2	97.6	91.3	64.2	12.4	89.3	64.7	34.5	44.8
Mali	56.2	9.8	57.8	32.4	47.3	83.6	71.7	60.8	25.4	95.4	36.8	4.3	31.4
Mozambique	37.2	*	76.7	38.1	66.9	91.1	81.5	74.6	24.4	45.6	55.0	17.9	46.3
Namibia	9.1	19.3	71.2	48.5	65.9	94.2	89.5	83.6	*	75.7	71.6	3.8	43.0
Nigeria	27.3	11.2	33.2	17.4	52.8	51.2	42.1	41.3	5.7	*	33.7	6.0	19.9
Rwanda	*	*	80.5	87.3	82.4	98.9	95.2	86.4	*	99.7	27.5	11.2	80.1
Senegal	91.6	27.4	29.4	36.4	81.6	94.1	80.6	77.5	1.8	62.9	20.5	1.5	65.3
Sierra Leone	64.8	22.4	53.8	32.0	90.0	95.6	78.6	83.2	36.1	79.6	85.1	37.2	57.6
Tanzania	69.4	39.0	51.2	59.2	88.0	96.0	86.0	41.3	2.0	78.9	44.8	43.4	37.6
Togo	76.8	3.9	60.6	57.5	77.0	95.3	74.3	81.7	25.6	82.4	18.5	9.1	47.9
Uganda	78.5	53.2	66.1	65.5	80.6	96.3	80.0	61.6	6.5	99.6.	46.7	62.8	60.7
Zambia	92.1	1.0	65.8	72.5	81.9	94.9	84.9	76.5	7.4	95.3	64.1	36.0	59.8
Zimbabwe	*	55.0	57.6	47.7	54.3	89.9	81.9	67.4	*	94.6	40.5	*	18.0
Mean	52.8	23.0	52.9	43.3	73.1	88.4	76.5	59.8	14.2	85.8	39.5	20.9	42.4

1. Fansidar; 2. Postnatal Care; 3. Immediate Breastfeeding; 4. Exclusive Breastfeeding; 5. Deworming
* No data

Table 3.6: Posterior estimates obtained from the bivariate geostatistical variable selection. Results are presented only for the statistically important associations between health interventions and U5M.

	# count	Angola	Benin	Burkina Faso	Burundi	Cameroon	Chad	DRC
H_1ITN	3							
H_1ITN2	1							
P_ITNA	3							
P_ITNS	4							
P_ITN5	6							-0.36 (-0.74,-0.01)
P_ITNU	3							-0.61 (-1.13,-0.09)
Improved Water	3			-0.47 (-0.74,-0.20)				
Improved Sanitation	12		-1.00 (-1.57,-0.42)	-0.74 (-1.19,-0.29)	-0.70 (-1.07,-0.34)		-0.69 (-1.14,-0.24)	-0.53 (-0.85,-0.20)
ANC provider	26	-0.78 (-1.33,-0.22)	-0.90 (-1.39,-0.42)	-0.68 (-1.01,-0.34)	-2.22 (-3.25,-1.19)	-1.12 (-1.63,-0.61)	-0.43 (-0.78,-0.08)	
4+ ANC visits	19	-0.66 (-1.22,-0.11)	-0.96 (-1.47,-0.46)	-0.53 (-1.05,-0.01)		-1.30 (-1.78,-0.82)		
Fansidar	12	-0.61 (-1.18,-0.04)	-1.10 (-1.58,-0.62)				-0.59 (-1.05,-0.14)	
Post. Care	4			-0.65 (-1.21,-0.09)				
Immediate	14					-0.38 (-0.77,0)		
Exclusive	2							
Tetanus	21	-0.78 (-1.35,-0.21)	-1.13 (-1.62,-0.64)	-1.80 (-2.39,-1.21)		-1.30 (-1.86,-0.75)	-0.42 (-0.79,-0.06)	
BCG	28	-0.82 (-1.38,-0.26)	-1.94 (-2.51,-1.38)	-3.09 (-3.52,-2.66)	-3.91 (-4.77,-3.04)	-2.85 (-3.29,-2.42)	-0.58 (-0.83,-0.33)	-1.24 (-1.54,-0.93)
DPT	28	-1.36 (-2.12,-0.60)	-1.61 (-2.15,-1.08)	-2.02 (-2.40,-1.63)	-3.56 (-4.40,-2.72)	-1.64 (-2.00,-1.30)	-0.61 (-0.93,-0.28)	-0.97 (-1.27,-0.67)
Polio	25	-1.07 (-1.88,-0.27)	-1.43 (-2.00,-0.87)	-1.77 (-2.13,-1.41)	-2.47 (-3.20,-1.74)	-1.30 (-1.67,-0.93)	-0.60 (-1.00,-0.21)	-0.66 (-1.07,-0.25)
Measles	28	-0.94 (-1.66,-0.22)	-1.52 (-2.15,-0.89)	-2.39 (-2.81,-1.96)	-2.87 (-3.67,-2.08)	-2.19 (-2.61,-1.76)	-0.71 (-1.01,-0.42)	-1.40 (-1.77,-1.04)
Vitamin A	27		-0.74 (-1.09,-0.39)	-1.00 (-1.31,-0.70)	-1.55 (-2.02,-1.08)	-0.88 (-1.32,-0.45)	-0.56 (-0.89,-0.22)	-0.86 (-1.18,-0.54)
Iron	8		-0.48 (-0.93,-0.02)			-1.40 (-2.38,-0.42)		-1.07 (-1.60,-0.55)
Iodized Salt	4		-0.68 (-1.32,-0.05)					
ORS	4					-1.80 (-3.48,-0.12)		
ACT	2						-0.90 (-1.66,-0.15)	
Deworming	26		-0.83 (-1.27,-0.39)	-0.50 (-1.00,-0.01)	-1.59 (-2.11,-1.08)	-1.72 (-2.17,-1.28)	-0.37 (-0.73,-0.01)	-0.89 (-1.25,-0.54)

count: sum of the number of countries, for which an intervention was statistically associated with under-five mortality.

	Egypt	Ethiopia	Gabon	Ghana	Guinea	Ivory Coast	Kenya
H_1ITN					-0.66 (-1.20,-0.13)		
H_1ITN2							
P_ITNA					-0.80 (-1.58,-0.02)		
P_ITNS							
P_ITN5				-0.73 (-1.42,-0.03)			
P_ITNU							
Improved Water							
Improved Sanitation	-0.59 (-1.10,-0.08)	-1.11 (-1.84,-0.37)			-0.89 (-1.40,-0.39)	-0.43 (-0.85,-0.01)	
ANC provider	-0.81 (-1.48,-0.13)	-0.97 (-1.38,-0.56)	-1.62 (-2.47,-0.76)	-2.42 (-3.57,-1.27)	-0.99 (-1.43,-0.54)	-1.50 (-2.11,-0.89)	-1.11 (-1.67,-0.56)
4+ ANC visits		-1.29 (-1.79,-0.79)		-1.27 (-2.11,-0.42)	-1.08 (-1.61,-0.55)	-0.88 (-1.39,-0.36)	
Fansidar				-1.87 (-2.81,-0.94)	-0.61 (-1.20,-0.03)		
Post. Care				-1.00 (-1.96,-0.04)			
Immediate		-0.58 (-1.03,-0.13)	-0.81 (-1.39,-0.23)	-0.65 (-1.28,-0.03)		-0.46 (-0.92,-0.01)	-0.58 (-1.04,-0.11)
Exclusive							
Tetanus	-0.96 (-1.51,-0.41)	-1.15 (-1.62,-0.68)	-1.69 (-2.60,-0.78)	-1.72 (-2.74,-0.69)	-1.30 (-1.84,-0.76)	-1.39 (-1.96,-0.83)	
BCG	-5.90 (-6.29,-5.52)	-1.49 (-2.02,-0.96)	-2.32 (-3.02,-1.61)	-4.80 (-5.85,-3.76)	-1.75 (-2.19,-1.30)	-1.78 (-2.23,-1.33)	-5.60 (-6.23,-4.98)
DPT	-4.55 (-5.05,-4.06)	-1.36 (-1.91,-0.82)	-1.11 (-1.98,-0.24)	-2.43 (-3.29,-1.58)	-0.93 (-1.36,-0.49)	-1.37 (-1.77,-0.97)	-2.86 (-3.32,-2.39)
Polio	-3.77 (-4.29,-3.24)	-2.13 (-2.93,-1.33)		-1.33 (-2.04,-0.64)	-1.13 (-1.61,-0.66)	-1.17 (-1.58,-0.76)	-0.81 (-1.15,-0.48)
Measles	-1.99 (-2.54,-1.44)	-1.57 (-2.21,-0.92)	-1.48 (-2.34,-0.63)	-3.61 (-4.62,-2.60)	-1.33 (-1.80,-0.87)	-1.51 (-2.04,-0.97)	-2.61 (-3.12,-2.10)
Vitamin A	-1.33 (-2.10,-0.56)	-1.20 (-1.66,-0.75)	-1.23 (-2.04,-0.41)	-1.13 (-1.83,-0.44)	-0.79 (-1.25,-0.32)	-0.68 (-1.12,-0.24)	-1.28 (-1.68,-0.88)
Iron	-1.60 (-2.77,-0.44)						
Iodized Salt							
ORS					-1.78 (-3.42,-0.14)		
ACT				-0.61 (-1.02,-0.19)			
Deworming	-1.21 (-2.25,-0.18)	-1.77 (-2.62,-0.91)	-1.41 (-2.22,-0.60)	-1.43 (-2.29,-0.59)	-1.65 (-2.34,-0.95)	-1.02 (-1.53,-0.51)	-1.52 (-2.15,-0.88)

	Liberia	Malawi	Mali	Mozambique	Namibia	Nigeria	Rwanda
H_1ITN							-0.77 (-1.53,-0.02)
H_1ITN2							
P_ITNA							-0.72 (-1.45,0)
P_ITNS		-0.59 (-1.06,-0.13)	-0.96 (-1.68,-0.25)				
P_ITN5		-0.72 (-1.12,-0.33)	-0.93 (-1.53,-0.33)				
P_ITNU		-0.56 (-1.01,-0.12)					
Improved Water						-0.34 (-0.48,-0.20)	
Improved Sanitation						-0.41 (-0.61,-0.22)	
ANC provider	-1.26 (-2.07,-0.45)	-1.99 (-2.71,-1.27)	-1.23 (-1.72,-0.74)	-0.94 (-1.59,-0.28)	-1.45 (-2.52,-0.38)	-0.83 (-1.10,-0.56)	-1.99 (-3.36,-0.63)
4+ ANC visits	-0.78 (-1.38,-0.18)	-0.90 (-1.45,-0.35)	-1.51 (-2.13,-0.88)	-0.78 (-1.35,-0.21)	-1.01 (-1.85,-0.16)	-0.80 (-1.09,-0.50)	-1.50 (-2.38,-0.62)
Fansidar		-1.94 (-2.63,-1.25)		-0.98 (-1.54,-0.42)		-0.93 (-1.35,-0.51)	
Post. Care						-1.72 (-2.34,-1.09)	
Immediate		-0.75 (-1.23,-0.27)		-0.62 (-1.09,-0.14)	-0.90 (-1.66,-0.13)	-0.71 (-0.98,-0.43)	-1.45 (-2.33,-0.56)
Exclusive							
Tetanus	-1.15 (-1.82,-0.48)	-1.56 (-2.19,-0.93)	-1.08 (-1.64,-0.52)	-0.81 (-1.38,-0.24)		-0.88 (-1.16,-0.59)	
BCG	-1.94 (-2.46,-1.42)	-2.79 (-3.52,-2.06)	-1.40 (-1.82,-0.98)	-3.08 (-3.61,-2.55)	-5.89 (-7.01,-4.77)	-1.11 (-1.33,-0.89)	-8.26 (-9.56,-6.96)
DPT	-1.14 (-1.63,-0.65)	-2.18 (-2.89,-1.47)	-1.16 (-1.59,-0.73)	-1.83 (-2.28,-1.40)	-2.48 (-3.30,-1.66)	-1.30 (-1.56,-1.05)	-5.44 (-6.43,-4.44)
Polio	-0.79 (-1.32,-0.26)	-0.99 (-1.59,-0.39)	-0.96 (-1.52,-0.41)	-1.44 (-1.85,-1.02)	-1.59 (-2.36,-0.81)	-1.40 (-1.73,-1.07)	-2.81 (-3.53,-2.09)
Measles	-1.83 (-2.40,-1.26)	-1.30 (-2.02,-0.58)	-1.30 (-1.77,-0.83)	-2.59 (-3.09,-2.09)	-2.69 (-3.54,-1.83)	-1.25 (-1.51,-0.99)	-0.74 (-1.34,-0.14)
Vitamin A	-0.75 (-1.20,-0.30)	-1.53 (-2.00,-1.06)	-1.21 (-1.63,-0.79)	-2.00 (-2.48,-1.53)	-2.83 (-3.58,-2.08)	-1.11 (-1.38,-0.84)	-3.28 (-4.14,-2.42)
Iron	-0.70 (-1.39,-0.01)		-1.46 (-2.06,-0.86)	-0.72 (-1.27,-0.17)			
Iodized Salt		-0.40 (-0.71,-0.09)					
ORS				-1.44 (-2.72,-0.15)			
ACT							
Deworming	-0.78 (-1.33,-0.22)	-1.48 (-1.97,-1.00)	-0.91 (-1.43,-0.40)	-1.26 (-1.75,-0.78)	-1.07 (-1.77,-0.37)	-1.05 (-1.42,-0.68)	-3.20 (-4.07,-2.32)

	Senegal	Sierra Leone	Tanzania	Togo	Uganda	Zambia	Zimbabwe
H_1ITN					-0.95 (-1.45,-0.46)		
H_1ITN2					-0.70 (-1.15,-0.25)		
P_ITNA					-0.89 (-1.37,-0.42)		
P_ITNS					-0.69 (-1.15,-0.22)	-0.60 (-1.13,-0.06)	
P_ITN5					-0.59 (-1.01,-0.18)	-0.51 (-0.94,-0.08)	
P_ITNU	-0.69 (-1.28,-0.10)						
Improved Water				-0.42 (-0.74,-0.09)			
Improved Sanitation	-0.83 (-1.32,-0.34)			-1.66 (-2.69,-0.64)			
ANC provider	-1.84 (-2.96,-0.72)	-1.02 (-1.91,-0.12)		-0.83 (-1.30,-0.37)	-1.40 (-2.19,-0.60)	-1.05 (-1.78,-0.33)	-2.38 (-3.10,-1.66)
4+ ANC visits	-2.35 (-3.26,-1.45)	-0.76 (-1.29,-0.22)		-1.09 (-1.76,-0.42)	-0.78 (-1.38,-0.18)		-1.81 (-2.60,-1.03)
Fansidar	-1.91 (-2.93,-0.89)			-1.60 (-2.22,-0.97)	-1.04 (-1.68,-0.40)	-0.92 (-1.57,-0.28)	
Post. Care	-1.91 (-2.65,-1.18)						
Immediate				-0.64 (-1.08,-0.20)		-0.48 (-0.91,-0.05)	-0.77 (-1.49,-0.06)
Exclusive						-1.76 (-3.03,-0.49)	-1.56 (-2.41,-0.70)
Tetanus	-1.69 (-2.76,-0.61)	-1.20 (-2.05,-0.35)		-1.47 (-2.23,-0.71)	-1.12 (-1.79,-0.46)		-2.23 (-3.02,-1.45)
BCG	-2.33 (-3.20,-1.46)	-3.46 (-4.05,-2.86)	-1.93 (-2.77,-1.08)	-3.80 (-4.54,-3.07)	-2.76 (-3.54,-1.98)	-4.58 (-5.25,-3.91)	-1.68 (-2.61,-0.75)
DPT	-2.13 (-2.86,-1.40)	-1.46 (-1.90,-1.03)	-1.13 (-1.88,-0.38)	-1.97 (-2.52,-1.43)	-2.16 (-2.82,-1.50)	-2.32 (-2.84,-1.80)	-1.06 (-2.01,-0.10)
Polio	-1.38 (-2.03,-0.73)	-1.41 (-1.84,-0.98)		-1.60 (-2.21,-0.99)	-1.81 (-2.46,-1.16)	-0.66 (-1.12,-0.20)	
Measles	-2.13 (-2.90,-1.37)	-2.22 (-2.79,-1.66)	-1.15 (-1.99,-0.31)	-2.18 (-2.85,-1.51)	-1.78 (-2.51,-1.06)	-3.42 (-4.02,-2.81)	-1.82 (-2.87,-0.78)
Vitamin A	-2.23 (-3.14,-1.33)	-1.77 (-2.25,-1.30)	-0.85 (-1.42,-0.29)	-2.63 (-3.22,-2.03)	-0.96 (-1.39,-0.52)	-2.05 (-2.55,-1.54)	-1.37 (-2.05,-0.69)
Iron		-0.90 (-1.39,-0.41)					
Iodized Salt		-0.40 (-0.72,-0.08)			-0.90 (-1.70,-0.11)		
ORS			-1.92 (-3.57,-0.28)				
ACT							
Deworming	-1.95 (-2.76,-1.15)	-0.94 (-1.45,-0.43)	-0.81 (-1.36,-0.27)	-0.68 (-1.19,-0.18)	-1.34 (-1.81,-0.87)	-1.66 (-2.16,-1.15)	

Figure 3.3: Spatially varying effects of percentage of households with at least one ITN on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

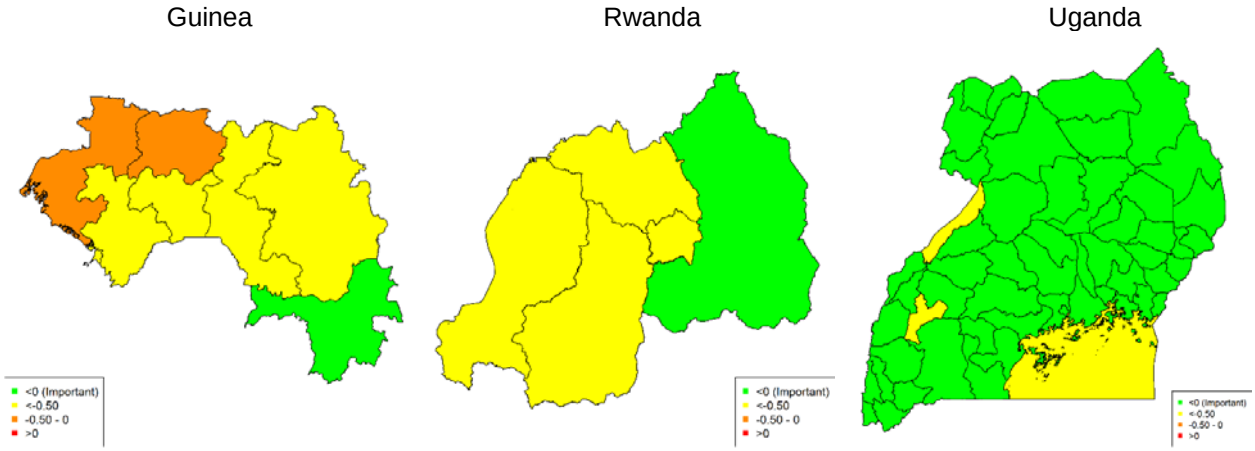


Figure 3.4: Spatially varying effects of percentage of households with at least one ITN for every two people on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

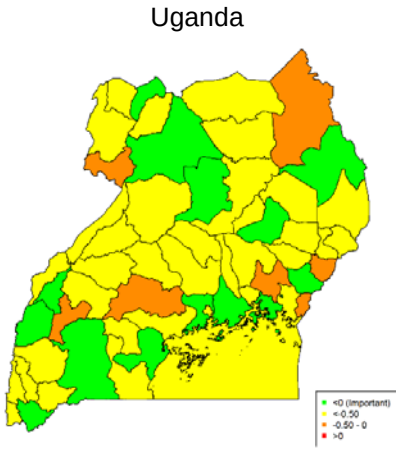


Figure 3.5: Spatially varying effects of percentage of population with access to an ITN within their household on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

Figure 3.6: Spatially varying effects of percentage of population in a household that slept under an ITN the previous night before the survey on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

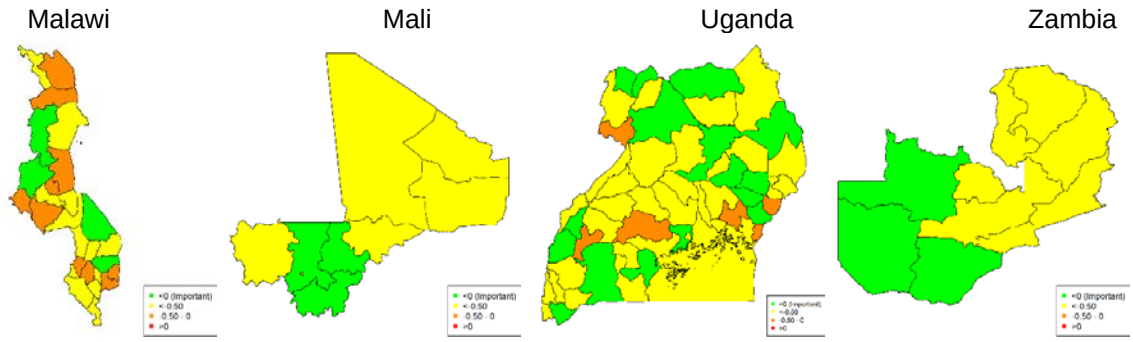


Figure 3.7: Spatially varying effects of percentage of existing ITNs used by the population in a household the previous night of the survey on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

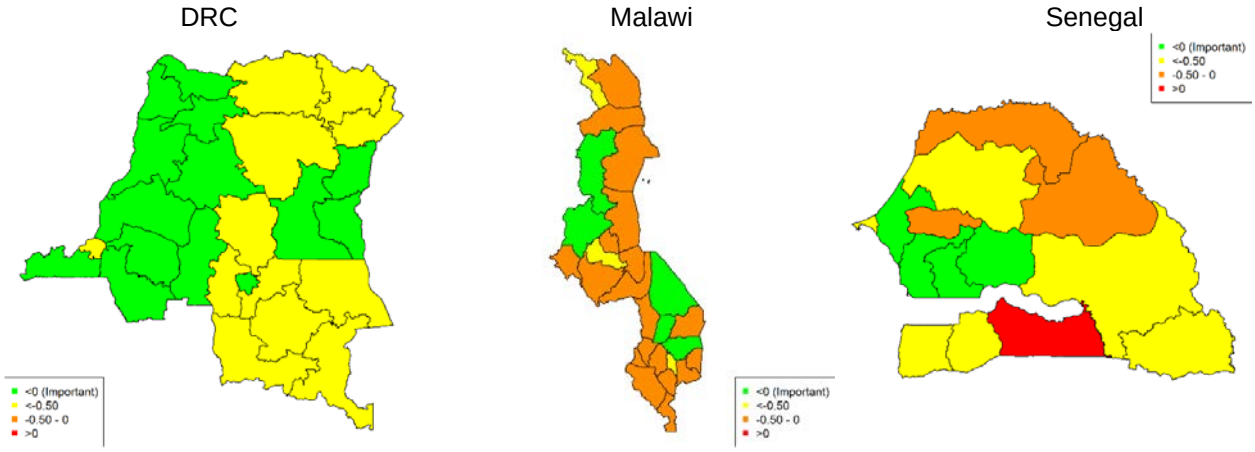


Figure 3.8: Spatially varying effects of percentage of households with improved source of drinking water on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

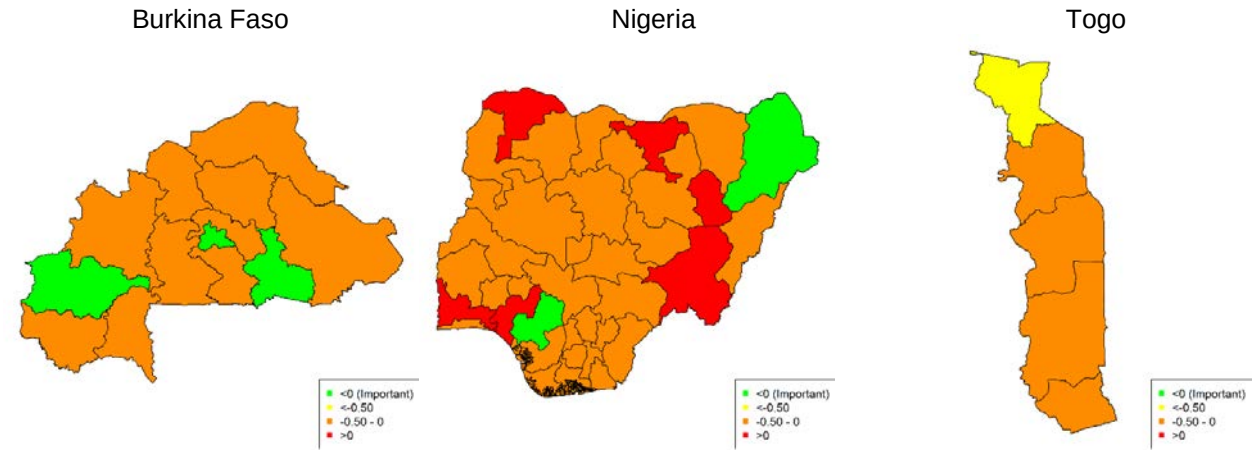


Figure 3.9: Spatially varying effects of percentage of percentage of new-borns receiving first postnatal check-up from a skilled provider within hours after delivery on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

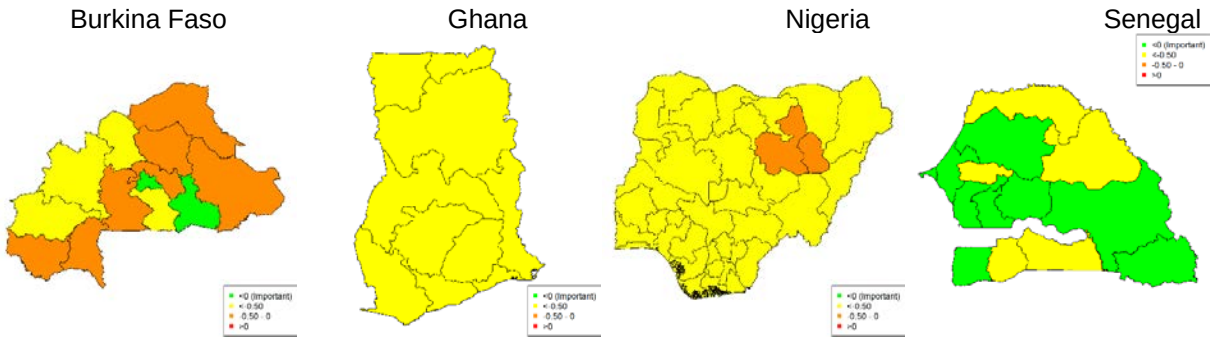


Figure 3.10: Spatially varying effects of percentage of infants which exclusively breastfed during the first six months after delivery on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

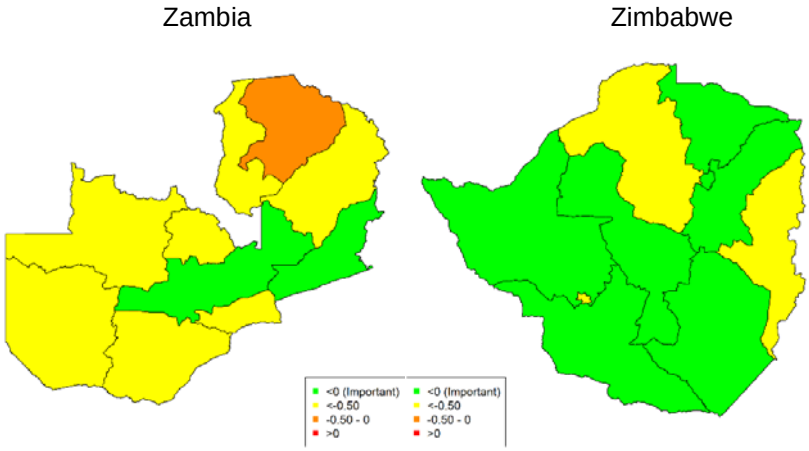


Figure 3.11: Spatially varying effects of percentage of children which lived in a household with positive test for iodized salt on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

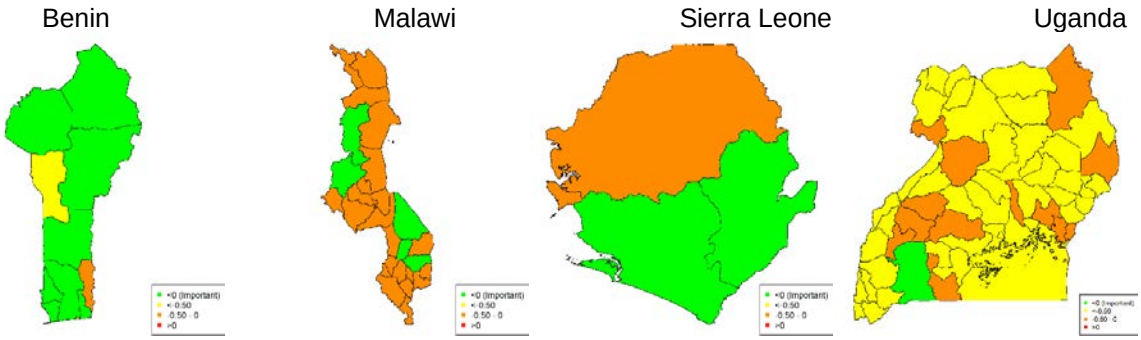


Figure 3.12: Spatially varying effects of percentage of children with diarrhoea given fluid from oral rehydration solution (ORS) on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

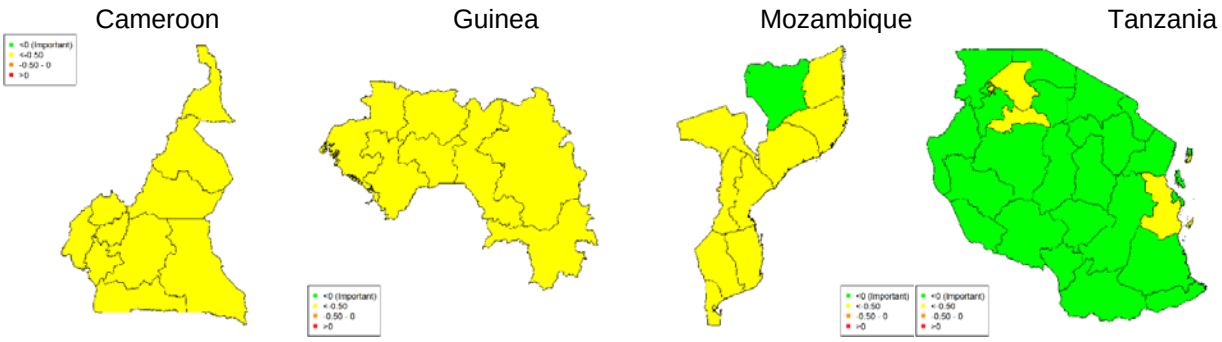


Figure 3.13: Spatially varying effects of percentage of children with fever during the two weeks prior to the survey which took artemisinin-combination therapy (ACT) on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

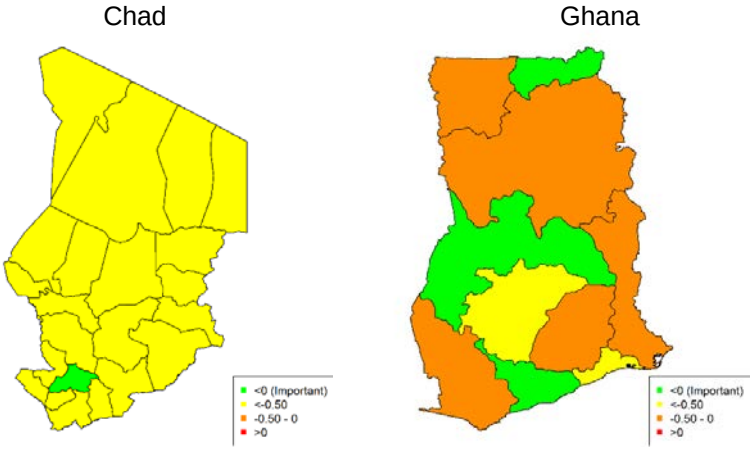


Figure 3.14: Spatially varying effects of percentage of under-five children in a household who slept under an ITN the night before the survey on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

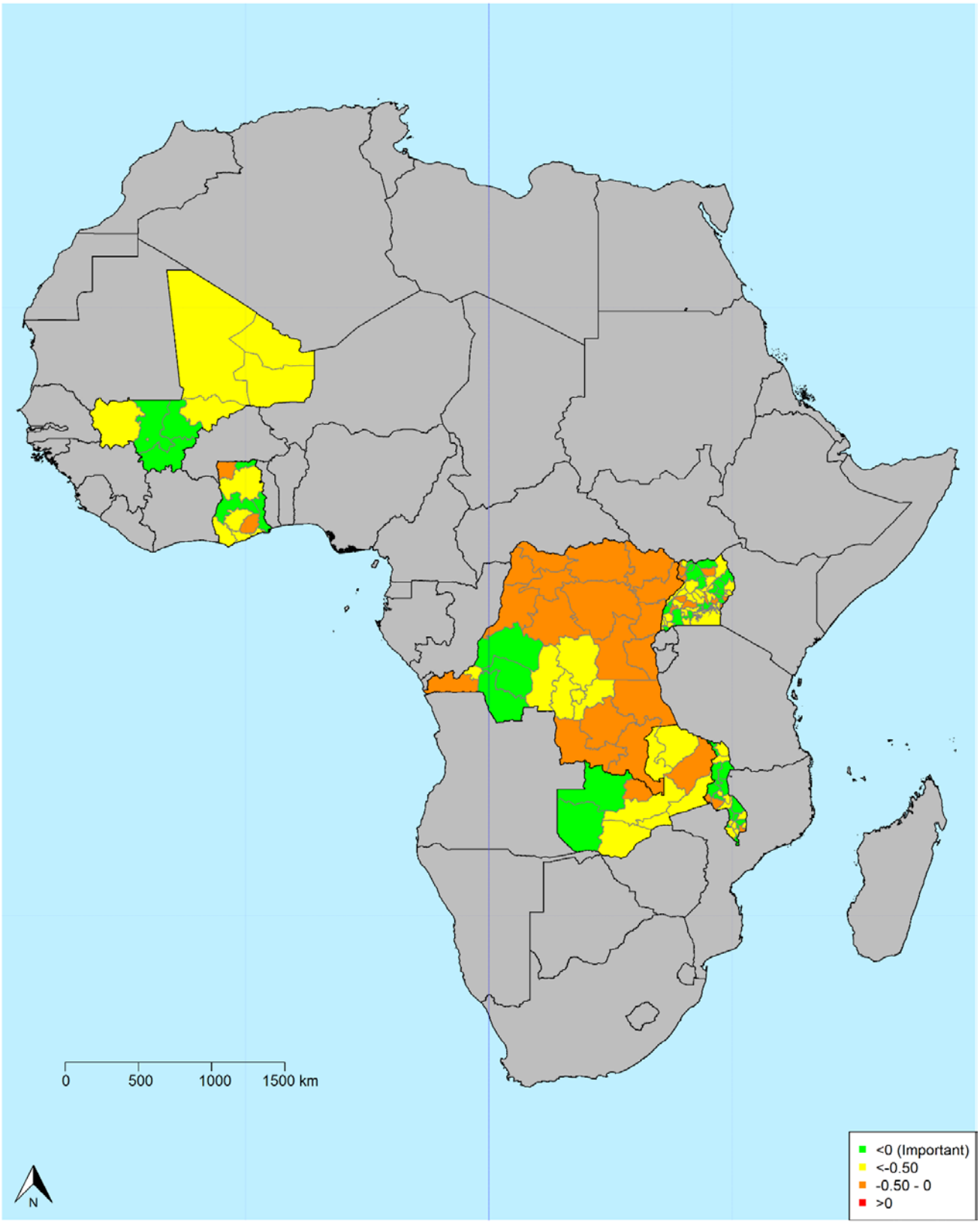


Figure 3.15: Spatially varying effects of percentage of households with improved sanitation facilities on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

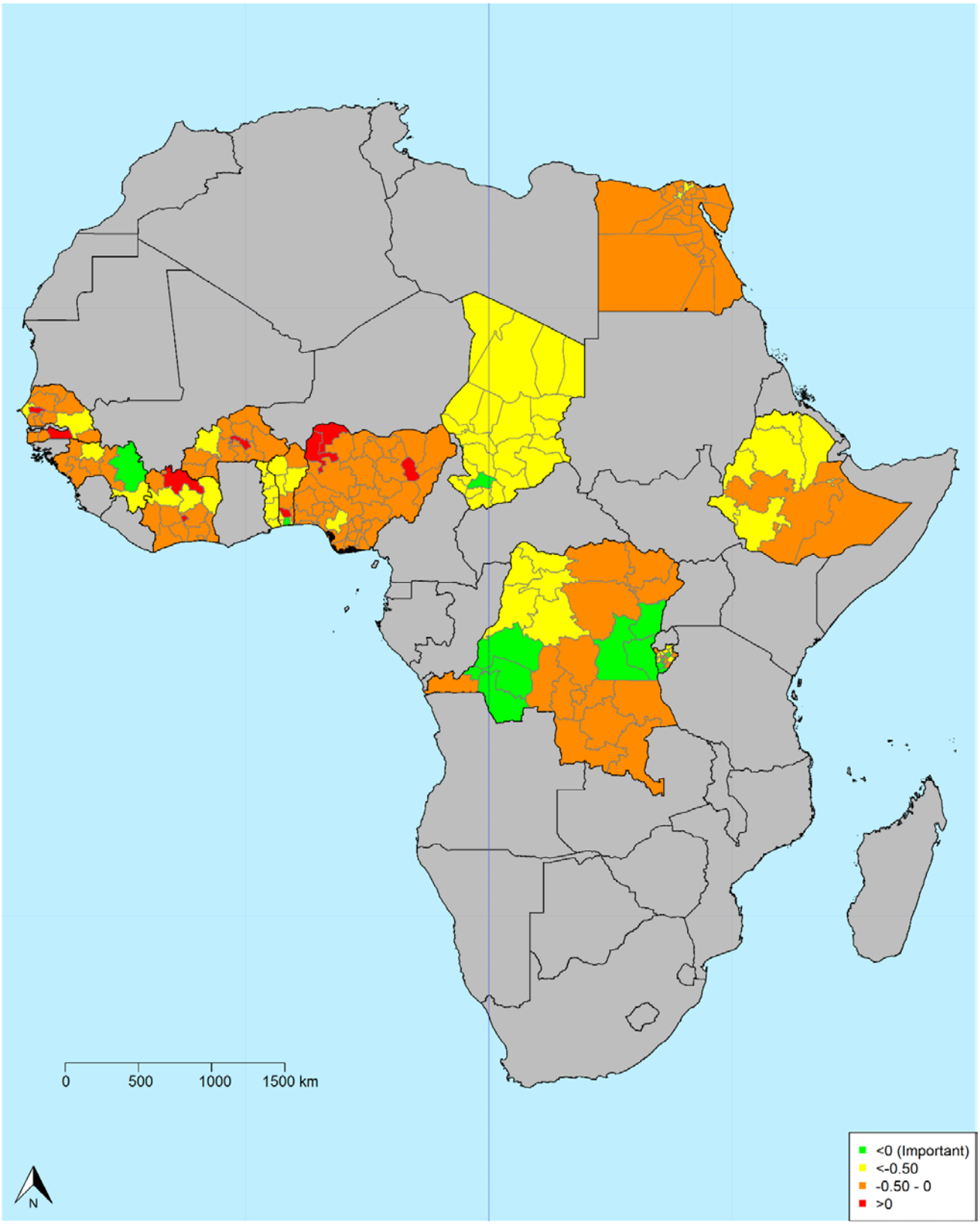


Figure 3.16: Spatially varying effects of percentage of pregnant mothers receiving ANC from a skilled provider on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

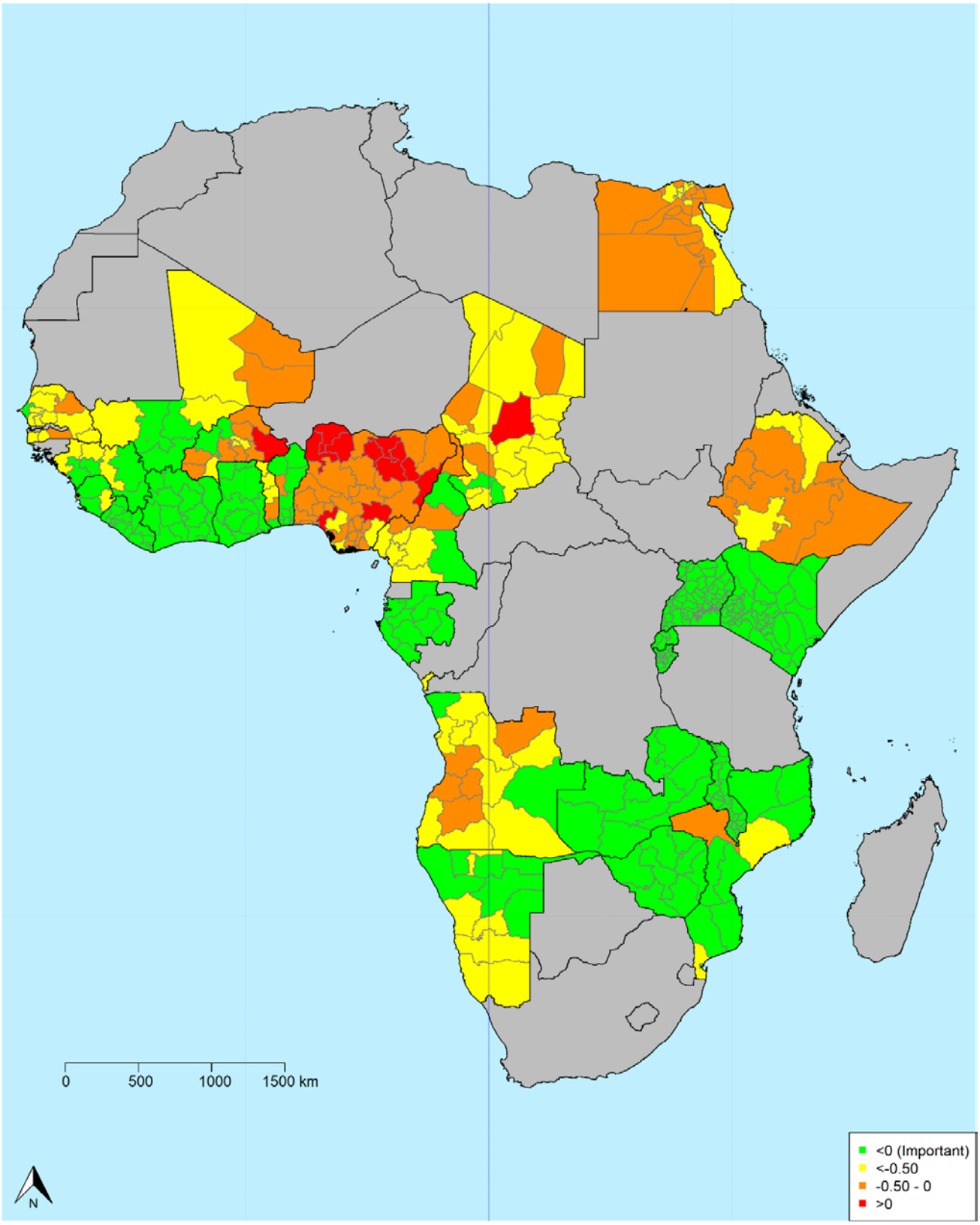


Figure 3.17: Spatially varying effects of percentage of pregnant woman making at least four ANC visits during their pregnancy on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

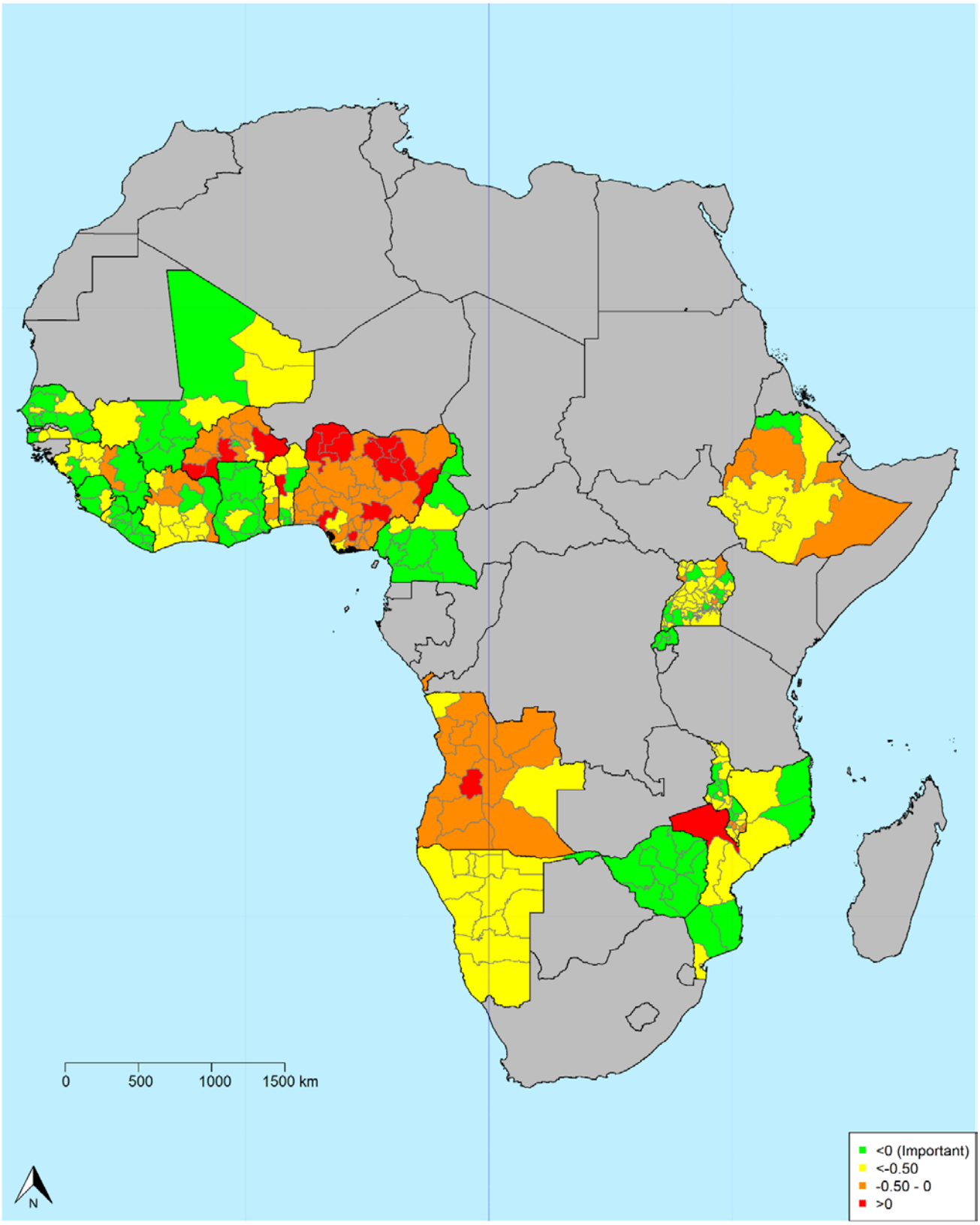


Figure 3.18: Spatially varying effects of percentage of percentage of pregnant woman which received Fansidar during pregnancy on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

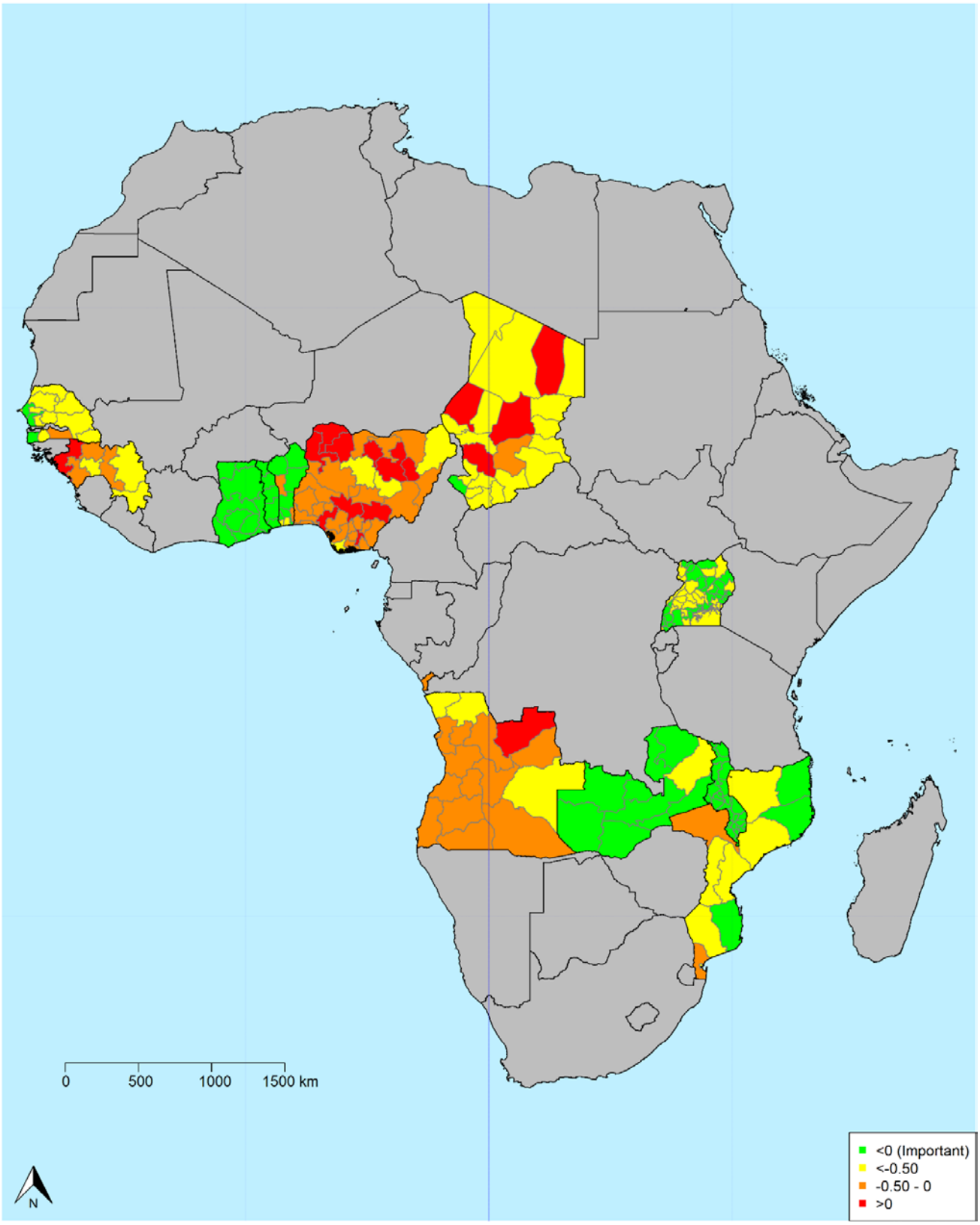


Figure 3.19: Spatially varying effects of percentage of infants which breastfed within one hour after birth on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

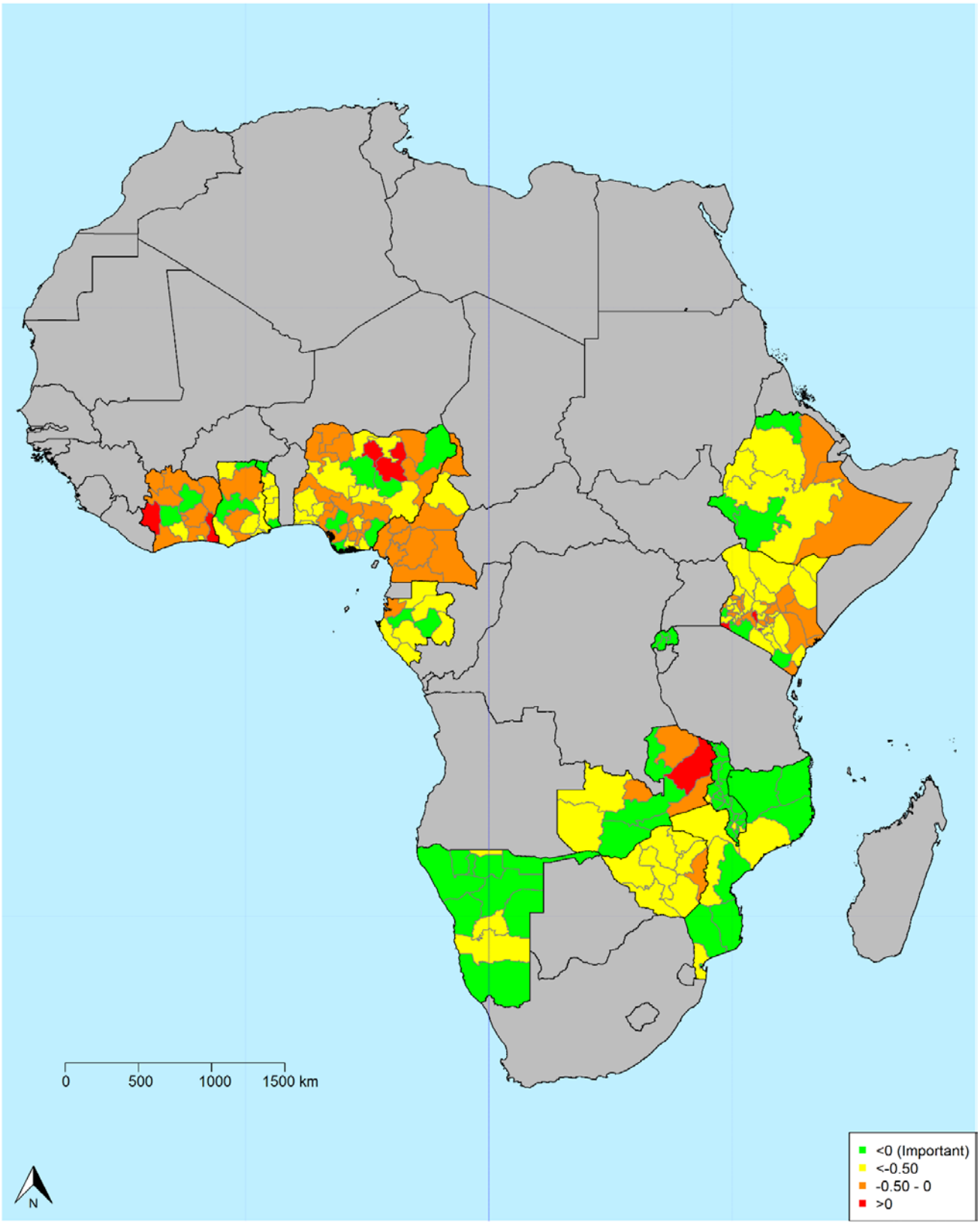


Figure 3.20: Spatially varying effects of percentage of last-born children protected against neonatal tetanus on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

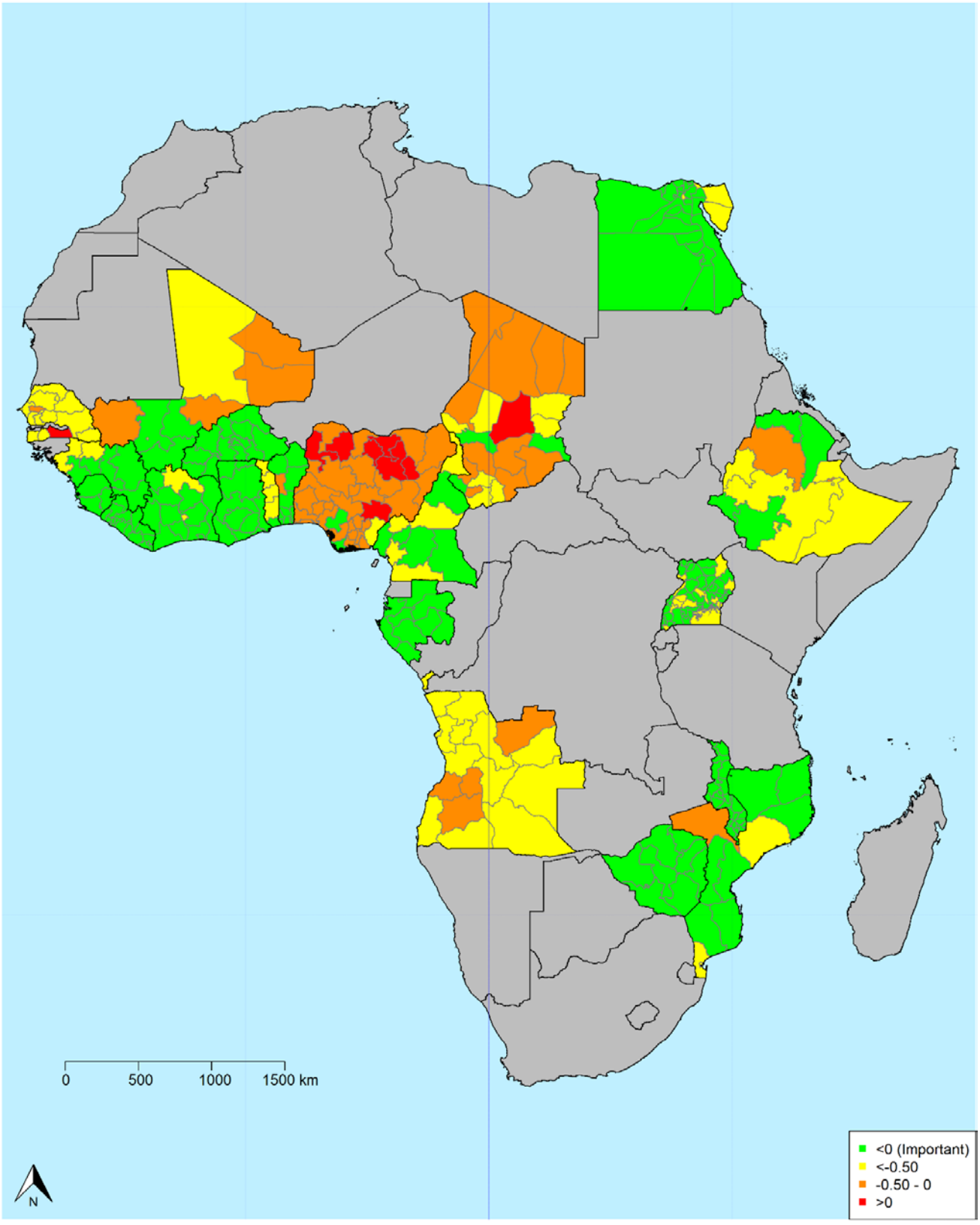


Figure 3.21: Spatially varying effects of percentage of children vaccinated against BCG on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

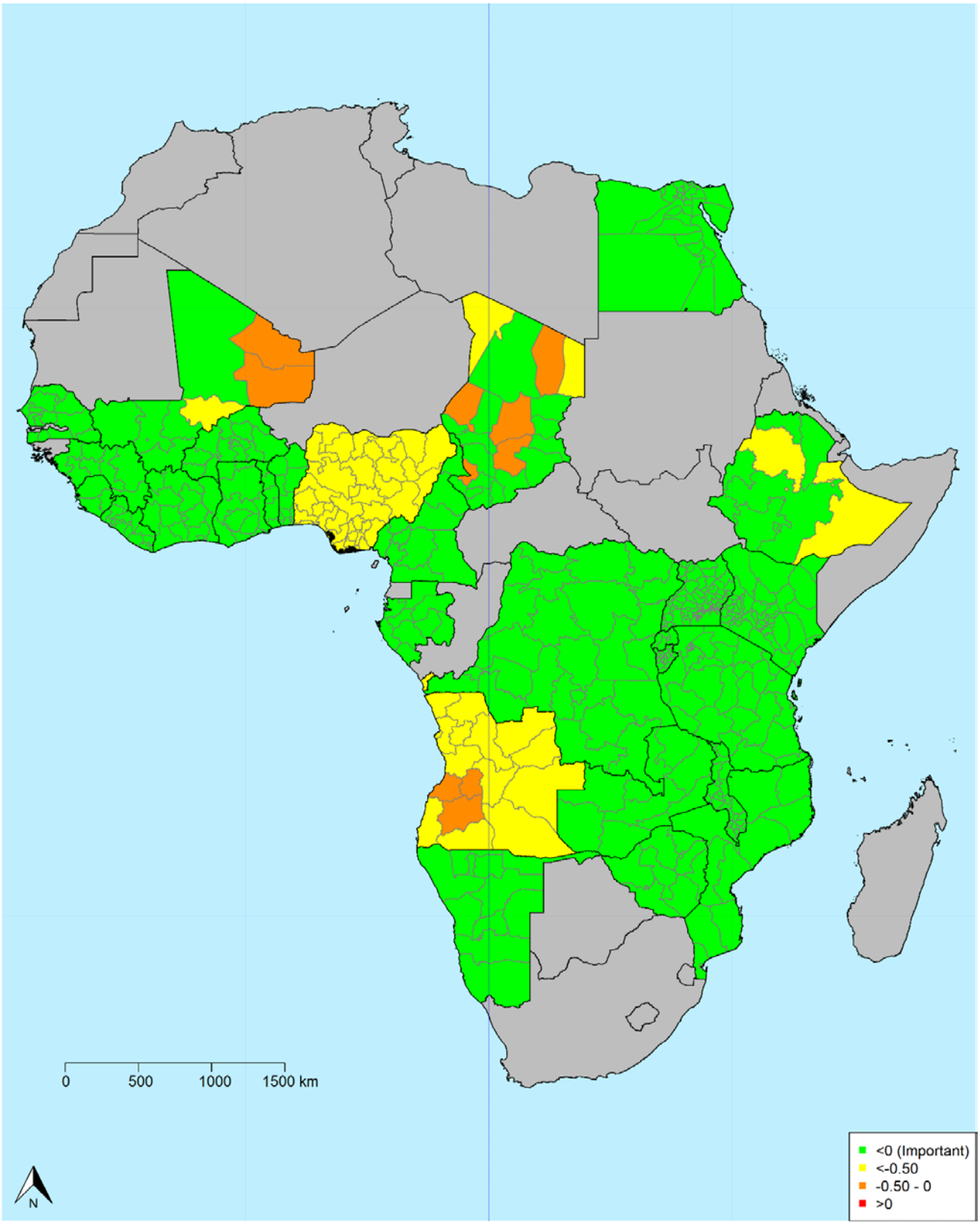


Figure 3.22: Spatially varying effects of percentage of children with complete DPT vaccination on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

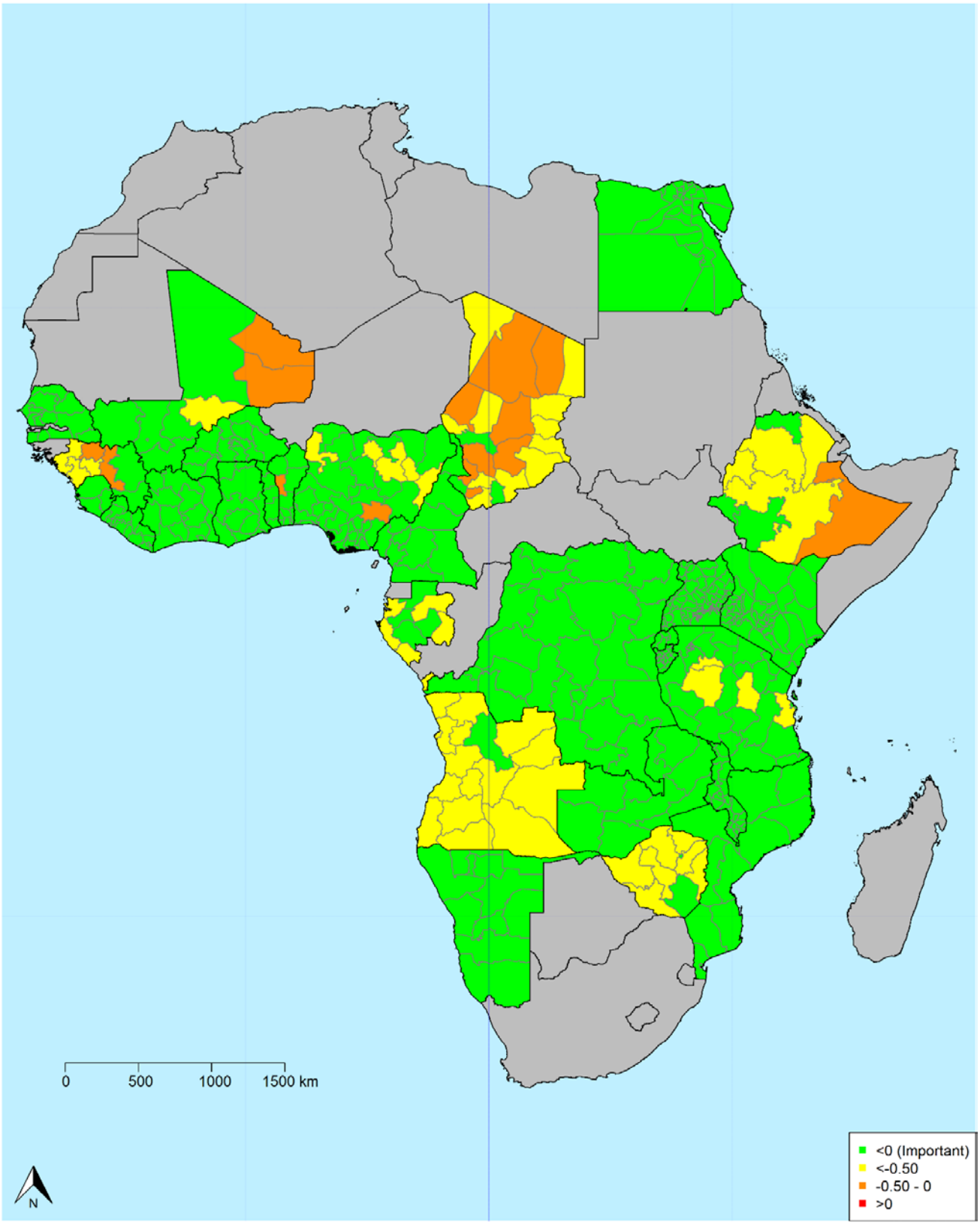


Figure 3.23: Spatially varying effects of percentage of children with complete Polio vaccination on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

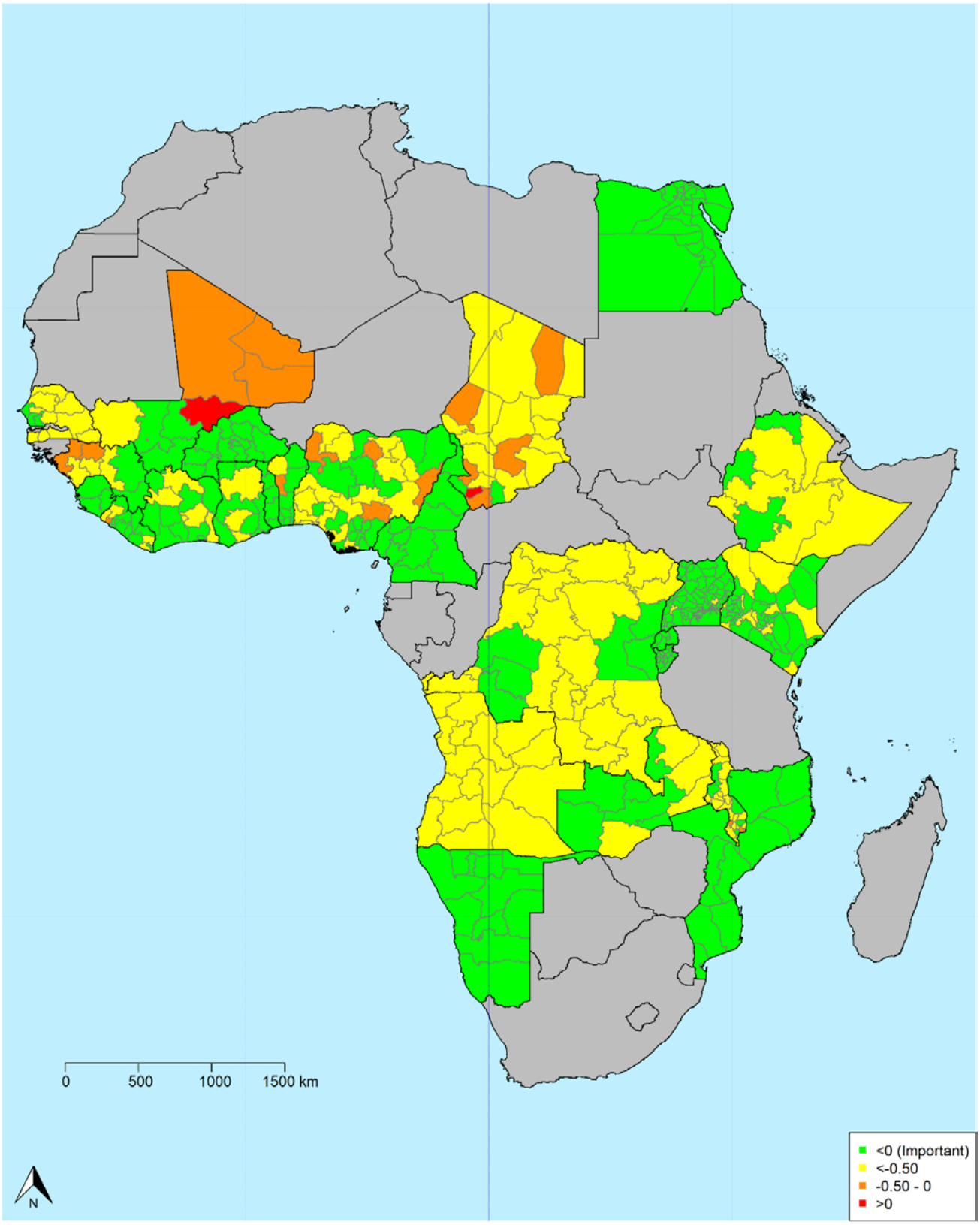


Figure 3.24: Spatially varying effects of percentage of children vaccinated against measles on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

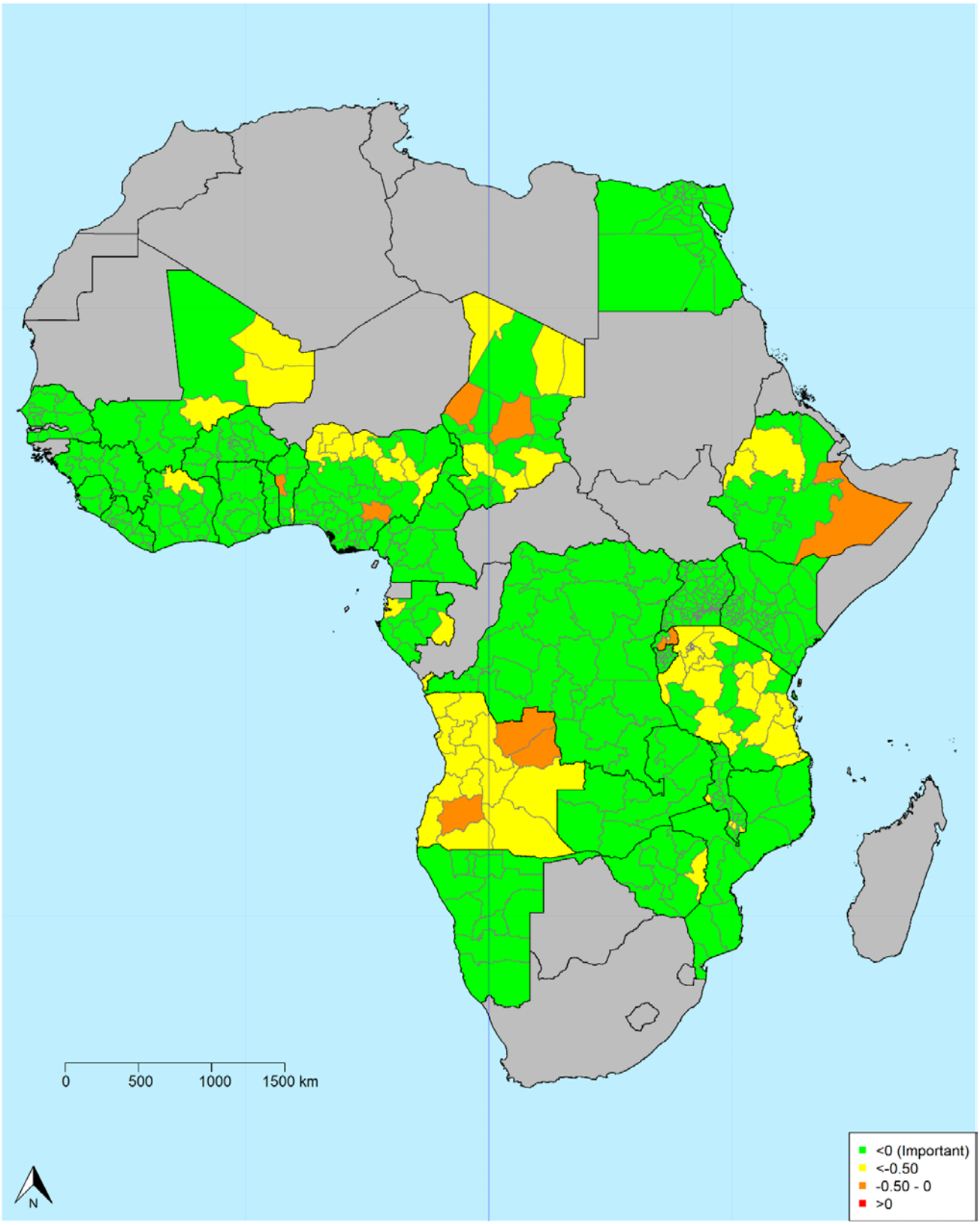


Figure 3.25: Spatially varying effects of percentage of children which received vitamin A supplements in the past 6 months on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

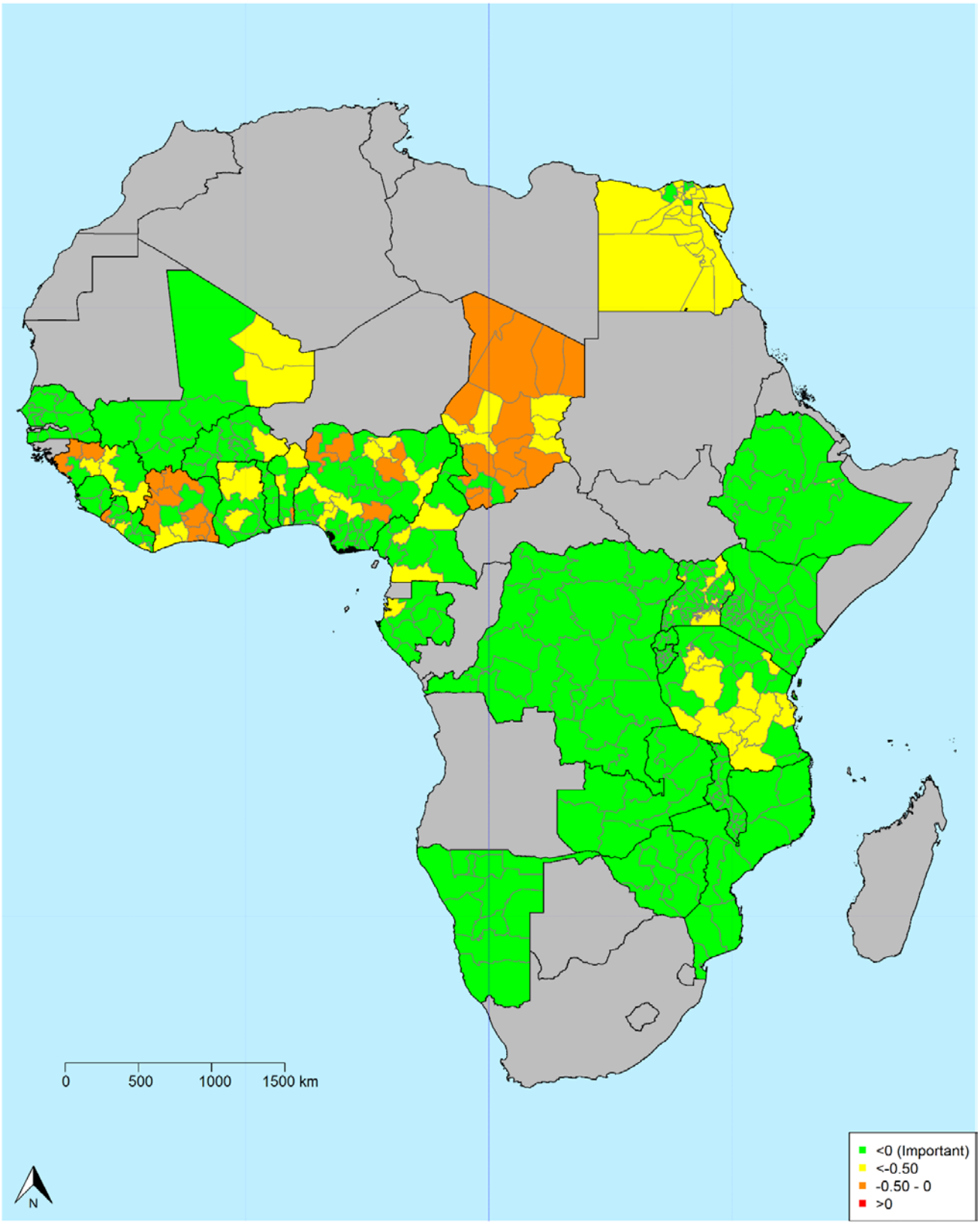


Figure 3.26: Spatially varying effects of percentage of children which received iron supplements in the past 7 days on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.

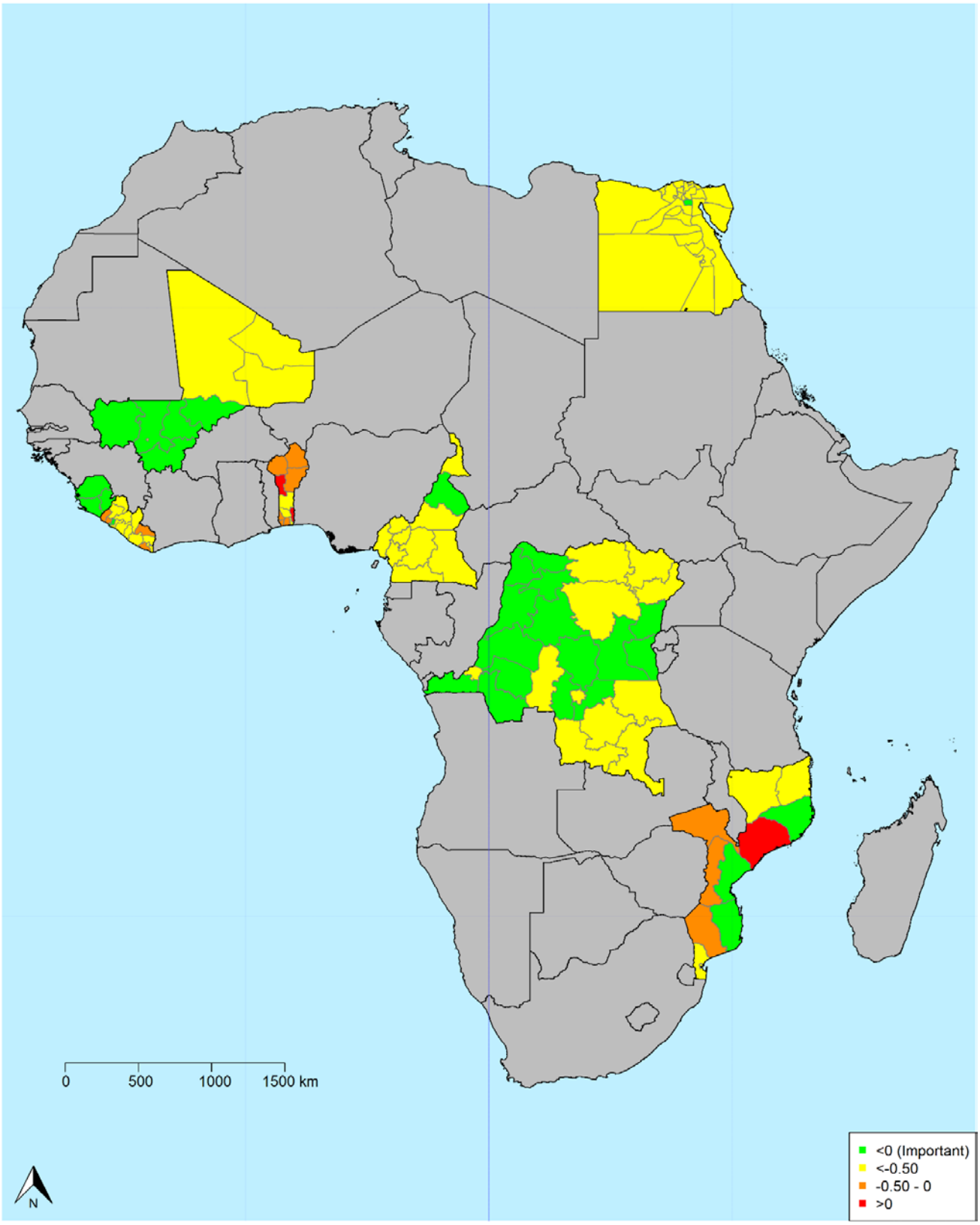
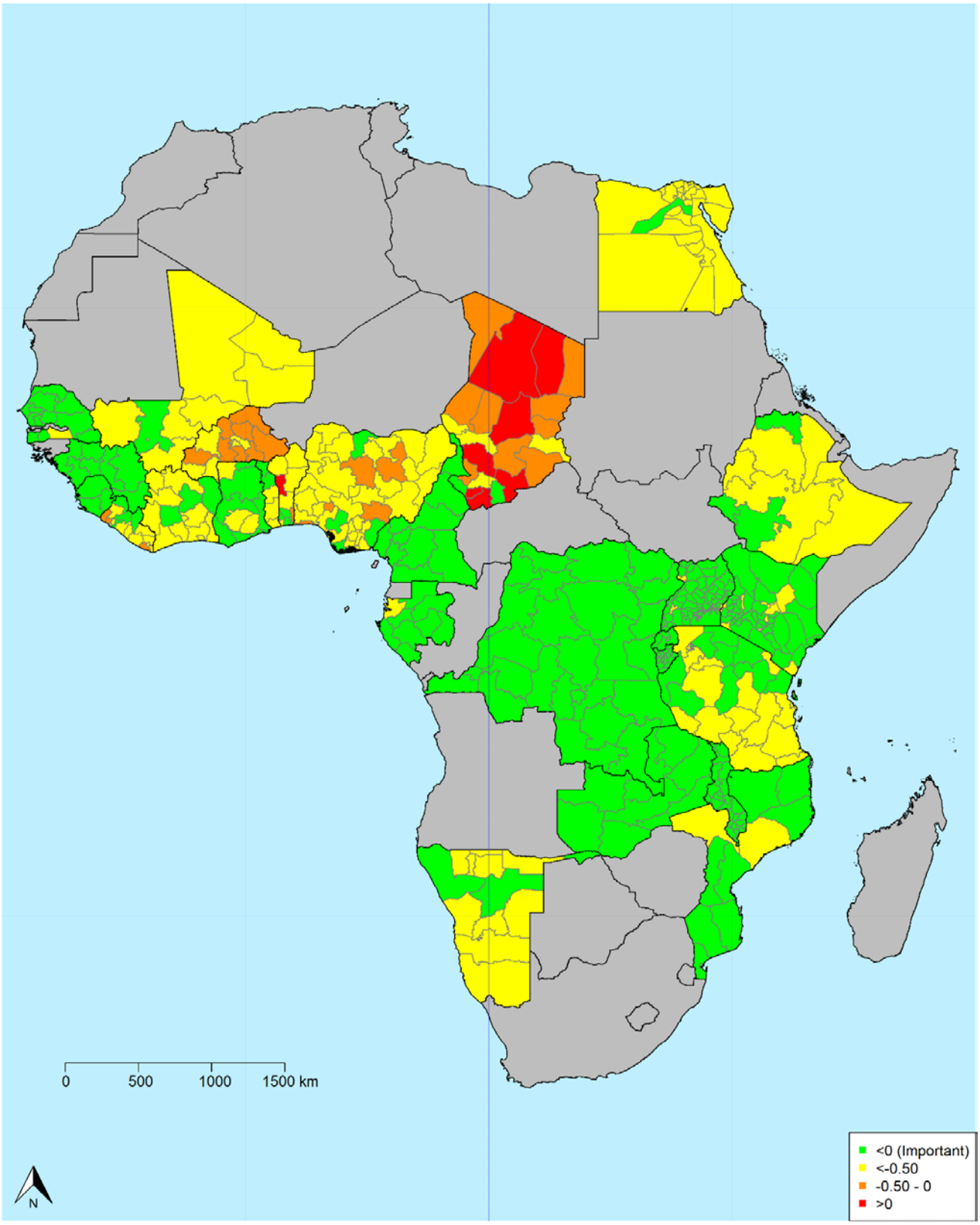


Figure 3.27: Spatially varying effects of percentage of children which took deworming medication in the past 6 months on under-five mortality at administrative level 1 estimated by a Bayesian geostatistical Weibull survival model.



Chapter 4

Geographical distribution of the association of childhood diseases with fever risk for under-five children in Africa

Isidoros Papaioannou[1,2], Jürg Utzinger[1,2], Penelope Vounatsou[1,2]

Affiliations:

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

Abstract

Febrile response is the clinical manifestation of various diseases and has traditionally being a frequent cause for medical consultation. In the sub-Saharan Africa (SSA) setting, development of fever is associated with all leading causes of mortality for children under the age of five, i.e. malaria, diarrhoea and acute respiratory infection (ARI). Crucially, presence of fever has been identified with malaria infection and so the clinical manifestation of fever leads to diagnosis and treatment of malaria without rigorous investigation of its aetiology. We utilized data coming from the Demographic and Health surveys, across 16 sub-Saharan Africa countries, in order to associate malaria parasitaemia, diarrhoea and ARI with the probability of developing fever. We employed a Bayesian geostatistical logistic regression model with spatially varying coefficients and quantified the associations at national and sub-national level. We developed province-level maps of the African continent in which we illustrated our estimates and calculated the Potential Attributable Fraction (PAF) in order to quantify the contribution of childhood diseases on fever. In Angola, Benin, Burkina Faso, Cameroon, Côte d'Ivoire, Ghana, Mozambique, Rwanda, Tanzania and Uganda, most fever cases were attributed to diarrhoea, as estimated by PAF, compared to malaria parasitaemia or ARI. For the rest six countries, i.e. Burundi, DRC, Guinea, Mali, Senegal and Togo, ARI was the main contributor. On average, the fever cases attributed to diarrhoea were 61.68%, to ARI 56.03% and to malaria parasitaemia 19.31%. We observed strong province-level disparities of the malaria-fever association, albeit almost everywhere, a homogeneous association between ARI and diarrhoea with fever. We concluded on the strong association of ARI and diarrhoea with fever and on the need for a rigorous investigation of the aetiology of febrile response upon suspected, or even confirmed, malaria cases.

4.1 Introduction

Febrile response is an outcome of inflammatory disease or infection, which involves the above-normal raising of the human body temperature and leads to improved protective processes for the immune system (Evans, 2015). Due to its presence as clinical manifestation of illness sine antiquity, fever constitutes one of the most frequent causes for medical consultation (Ogoina, 2011). The aetiology of fever can be attributed to an immense set of potential diseases as, for instance, in low and middle-income countries fever could be the outcome of blood parasites, e.g. malaria and trypanosomiasis, viral infections, e.g. dengue fever and Influenza, or bacterial zoonoses, e.g. q or trench fever (Prasad, 2015). Clinical manifestation of febrile response is particularly relevant for the age group of children less than five years, as pre-school children, elderly people and pregnant woman are the most vulnerable groups to contract diseases (WHO, 2019e). Geographically, the sub-Sahara African (SSA) countries are home to most under-five deaths globally and hence childhood fever is an important public health priority (WHO, 2019c). Specifically, febrile response is an outcome of the leading causes of under-five mortality in Africa, i.e. acute respiratory infection (ARI) manifested as pneumonia, diarrhoea and malaria, as well as of other infectious diseases (WHO. Regional office for Africa, 2019f).

Different data sources and methods have been considered to assess the aetiology of fever among under-five children in SSA. In Tanzania, a study recruited 1005 children with fever, below the age of 10, from two different outpatient clinics and after a systematic and thorough examination, it reported that 62.2% of the children had ARI, 13.3% had a non-malaria and non-typhoid infection and that only 10.5% of all children had malaria (D'Acremont, 2014). Additionally in Tanzania, a prospective cohort study monitored 870 hospitalizations due to fever across two sites in northern Tanzania and reported Chikungunya (10.2%) and Leptospirosis (7.7%) as the leading causes of fever among children less than 13 years old, compared to only 1.3% attributed to malaria (Crump, 2013). Yet, the prior study did not reported a cause for fever for the majority of the children, i.e. 64% and did not take into

account respiratory viruses. In Kenya, a case-control study for children under-five experiencing fever concluded that, among cases, 41.5% had a respiratory virus, 5.2% had malaria and 2.3% had group A streptococcal (GAS) pharyngitis, while for 37.1% an aetiology for febrile illness was not established (O'Meara, 2015). A recent narrative review focusing on the causes of febrile illness in SSA concluded that viral upper respiratory and self-limiting arboviral infections account for the majority, i.e. 60%, of children seeking treatment in health centres (Maze, 2018).

A common limitation shared among most studies investigating the contribution of illnesses to febrile response burden is the utilization of hospital-related data across a well-defined geographical area, i.e. specific, small in number, sites within a given country. Firstly, the selection of a very small number of sites, based on a non-systematic procedure, may not be representative of the dynamics of diseases at national level. Secondly, under-five children experiencing fever may not even attend a health facility for seeking diagnosis and treatment of fever. As depicted by a recent probabilistic modelling study, which analysed household survey data across 29 SSA countries, health-seeking behaviour for fever might be so low that, in an estimated 16 SSA countries, less than 50% of the under-five children with fever, which reside no further than two hours from a public health facility, would seek treatment (Alegana, 2018). Another shortage in the existing literature is that, to our knowledge, there is a scarcity of modelling studies trying to associate the contribution of diseases to fever at national level, but also exploring the geographical variation of the association within countries. An example of such a study for Uganda modelled the association of the leading causes of under-five mortality to fever, attributed the most cases to ARI and identified important regional variation of the effects (Nambuusi, 2019). Another issue which needs further investigation and that was also discussed in the aforementioned studies, is the over-diagnosis and hence over-treatment of malaria among under-five children with fever in SSA, which stresses the need for further investigation of the relation between childhood diseases and fever in malaria-endemic countries (Bisoffi, 2013). For instance, a recent modelling study considering 41 surveys across 21 SSA countries concluded that only 37.7% of all

malaria-positive, under-five children experiencing fever had developed a febrile response due to *Plasmodium falciparum* infection (Dalrymple, 2019).

Our work aims to assess the contribution of childhood diseases on febrile response among under-five children in SSA, by associating the prevalence of malaria parasitaemia, diarrhoea and ARI with fever. We utilized data coming from the cross-sectional, national-representative, household-based Demographic and Health surveys (DHS) for 16 SSA countries and employed a Bayesian hierarchical modelling approach. Our models incorporated spatial random effects to account for correlations in the fever outcome and a spatially varying coefficient component imposed to the disease exposure that aims to assess the geographical variation of the disease exposure on fever. Additionally, models were adjusted for maternal, socio-economic, demographic and individual-child characteristics. Finding of this study will provide further insights in the contribution of diseases to fever, depict geographical variations of these associations and contribute to the understanding of over-diagnosis and over-treatment of malaria in SSA.

4.2 Methods

Data sources

Our study is based on open-access surveys made available by the DHS program and remote sensing data. The DHS are cross-sectional surveys, typically conducted every five years, with a mandate to collect key health, demographic and nutrition related information. The surveys are represented on national and sub-national level, i.e. states, departments and residence, and are household-based. The goal of the DHS is to assist countries lacking strong vital registration systems to collect crucial information for evaluation and monitoring purposes. Questionnaires, biomarker tests and geographical information are some of the instruments used by the DHS to collect its data. DHS surveys are taking place in conjunction with the national agencies of each country. In order to protect the privacy of household members and respondents, the DHS program demands that before each interview or biomarker test, an informed consent statement is read to the respondent. It is therefore in

the respondent's judgment to accept or decline the invitation for participation in the interview or test. For children less than five years of age, the parent, guardian or caregiver is responsible for providing verbal informed consent. Verbal consent consists of a prescribed statement and the interviewer records the reply from the respondent. Additionally, the interviewer signs by name to having read the consent statement to the respondent. Written consent is not included (https://dhsprogram.com/publications/publication-dhsm7-dhs-questionnaires-and-manuals.cfm).

Across all 16 surveys, corresponding to the countries included in this study, we had available information on under-five children, which consisted of indicators on childhood diseases, maternal, socio-economic, demographic and individual child characteristics, health interventions, as well as the geographical coordinates of the spatial cluster within which each child was residing. Specifically, we had data on the presence of fever for a child during the two weeks preceding the survey, severe cough in the form of short, rapid breathing, diarrhoea over the two weeks preceding the survey and presence of malaria parasites that was confirmed from rapid diagnostic tests (RDTs). As the information on diseases was only available for the alive children, we aggregated the severe cough, diarrhoea and malaria parasitaemia at cluster level and considered that this corresponds to the cluster-level exposure that all children, within that cluster, experience. Presence of fever was retained at individual level. We assumed that the severe cough, expressed as cough accompanied with short, rapid breathing, corresponds to acute respiratory infection. Further, we had data on the sex of each child, the mode of delivery and its birth order, as well as the educational level of its mother and the age at which the mother gave birth. Based on the wealth quantile assigned to the household in which a child was residing, we associated each child with a wealth index. For similar reasons as with the childhood diseases, we aggregated curative or preventive health interventions at cluster level and assigned them to each child based on geographical information. The interventions considered were: improved source of drinking water and sanitation facilities; Tetanus, Bacillus Calmette–Guérin (BCG), Diphtheria, Tetanus, and Pertussis (DPT), polio and measles immunization; vitamin A and iron

supplementation; deworming medication; antenatal care (ANC) from a skilled provider during pregnancy; malaria bednets.

We downloaded remote sensing data from open access sources in order to adjust our models for other-than-malaria infectious diseases. We compiled information on land surface temperature at day, land surface temperature during night, distance to the nearest water bodies and the normalized difference vegetation index from the Moderate Resolution Imaging Spectroradiometer (MODIS) imaging sensor and the altitude of a location from the Shuttle Radar Topography Mission (SRTM). Also from MODIS, we extracted data on the land cover type, which we further transformed to three categories, i.e. forest, grassland and cropland, which we assigned to each child based on its geographical coordinates. From the U.S. Geological Survey-Earth Resources Observation Systems (USGSS) Data portal, we downloaded information on rainfall and from the Global Rural and Urban Mapping Project (GRUMP) we identified an either urban or rural residential status for each child.

Statistical analysis

We utilised Bayesian inference and developed a geostatistical logistic regression model, with a spatially varying coefficient component, in order to associate malaria parasitaemia, diarrhoea and ARI with fever across 16 SSA countries. The outcome of the model was a binary indicator that specified if an under-five child experienced fever or not and hence, we assumed that the outcome followed a Bernoulli distribution. The default logit-link function associated the probability of developing fever with the linear predictor, which incorporated an intercept, the exposure to a particular disease, confounders, a geostatistical intercept, which accounted for the spatial correlation in the fever outcome and a spatially varying part imposed to the disease exposure. Specifically, a Gaussian process with Matérn covariance matrix modelled the geostatistical intercept (Banerjee, 2014). The spatially varying coefficient accounted for the province-specific associations and was modelled via a multivariate Gaussian distribution with exchangeable structure. Before fitting the model, the Variance Inflation Factor was used to exclude from the linear predictor highly correlated

covariates and so to avoid collinearity (O'Brien, 2007). The posterior median from the regression coefficient of the disease exposure was interpreted as the national estimate of the disease-fever association. For the province-specific associations, we obtained a posterior sample of size 1,000 from the exchangeable structured effects, added them up with the national effect and summarized the estimates in terms of median and 95% credible interval. We applied the exponential function to all estimates in order to obtain the corresponding odds ratio. We summarized the national estimates in Table 5.2 and mapped the province-level associations in Figures 5.1–5.3. Throughout our work, we used terminology compatible with the Bayesian inference and called important coefficient what is called statistically significant from a frequentist perspective. Important coefficient means that the posterior distribution (95% Bayesian Credible Interval) does not include zero. We concluded our analysis by utilizing the Population Attributable Function (PAF) in order estimate the contribution of each childhood disease to fever (Leviton, 1973). PAF calculation for each of the provinces was based on the posterior sample of the spatially varying effects and so we summarized all PAF values in terms of posterior median and 95% Bayesian credible interval.

We used the RStudio (RStudio Team (2018), 2018), an integrated development environment of R software (R Core Team (2019), 2019), for our analysis, i.e. data management, statistical analysis, mapping efforts and PAF calculation. We used the Integrated Nested Laplace Approximation (Rue, 2009; Lindgren, 2011) (INLA) to conduct Bayesian inference. The Appendix contains further details concerning our analysis and model specification.

4.3 Results

The Burkina Faso DHS was the earliest survey included in our study, conducted in 2010 and accounted for the highest malaria parasitaemia prevalence (76.1%) and highest under-five morality rate (129 deaths per thousand live births) across all 16 DHS surveys included in our study. The Benin survey was the most recent, as it took place in 2017 and accounted for 540 unique survey locations, out of the total 7,741 sites incorporated in our study. Angola and

Uganda contributed the most to the total number of unique survey sites, while Senegal had the lowest number of locations, i.e. 214. Guinea, Mali and Côte d'Ivoire followed Burkina Faso as the countries with the highest malaria parasitaemia prevalence, ranging between 41% and 48%, while Rwanda and Senegal had the lowest. Burundi and Cameroon had the highest percentages of diarrhoea among their under-five children, accounting for more than 20%, while Uganda and Burundi the highest percentages for ARI. Mali was among the countries with the lowest diarrhoea, ARI and fever prevalence. Based on the raw data, Burundi, Uganda, Democratic Republic of the Congo (DRC) and Guinea were the countries that simultaneously had the highest ARI and fever prevalence. A descriptive analysis of all DHS surveys is presented in Table 5.1.

Across all countries the three main under-five diseases, i.e. malaria parasitaemia, diarrhoea and ARI, were increasing, in almost all occasions, the odds of a child developing fever. Only in Burkina Faso, Côte d'Ivoire, Mali, Senegal and Togo we did not find an important association between malaria parasitaemia and the probability of a febrile response. In Angola, Burkina Faso, Burundi, DRC, Ghana, Guinea, Mali, Senegal, Tanzania and Togo, the ARI had the highest odds ratio among all childhood diseases, with diarrhoea having the highest value for the rest of the countries. Across all cases, the highest odds ratio values were observed from the ARI-fever association, for instance in Senegal, Mali and Angola, while the odds ratio of malaria parasitaemia on fever were consistently the lowest from all childhood diseases for all countries. Results from the estimated odds ratios are presented in Table 5.2.

Our model-based maps, depicting the geographical distribution of the contribution of childhood diseases on fever, showed strong disparities for the malaria parasitaemia-fever association, albeit a homogeneous association of fever with diarrhoea and ARI. The regions of Kindia and Labé in Guinea, Sikasso in Mali, Comoé and Woroba in Côte d'Ivoire, Hauts-Bassins and Centre-Ouest in Burkina Faso, Upper-West in Ghana, Savanes and Plateaux in Togo and the Zou, Mono, Atlantique and Ouémé in Benin were the only provinces in West Africa with an important malaria parasitaemia-fever association. We observed more areas of

important association in Central and East Africa, i.e. several north and west departments of Cameroon, the Bas-Congo, Kivu and Bandundu provinces of DRC, northern Angola, several parts of Mozambique, most of Uganda, Rwanda and Burundi and all regions of Tanzania. Contrarily, we found almost everywhere an important association between diarrhoea or ARI prevalence with fever. Results from the estimated odds ratio for the childhood diseases on fever prevalence are presented at province-level in Figures 5.1–5.3.

In Angola, Benin, Burkina Faso, Cameroon, Côte d'Ivoire, Ghana, Mozambique, Rwanda, Tanzania and Uganda, most fever cases could be attributed to diarrhoea, as estimated by PAF, compared to malaria parasitaemia or ARI. For the rest six countries, i.e. Burundi, DRC, Guinea, Mali, Senegal and Togo, ARI was the main contributor, albeit in half of the countries the posterior estimates (median) were relatively close. On average, the fever cases attributed to diarrhoea were 61.68%, to ARI 56.03% and to malaria parasitaemia 19.31%. Across all countries, malaria parasitaemia had the lowest median PAF estimate. Strong variation on the contribution of childhood diseases to fever was observed at sub-national level as, for instance, in the Nord-Ouest province of Cameroon, approximately 25% of all fever cases can be attributed to diarrhoea, compared to a national average of 51.02% and an estimated contribution of 75.86% for the Nord province. PAF estimates are summarized in the Appendix.

4.4 Discussion

To our knowledge, this is the first modelling study associating the leading causes of under-five mortality with febrile response, using routinely collected, household-based survey data coming from the DHS for 16 SSA countries. We developed a Bayesian geostatistical logistic regression model, with spatially varying coefficients, in order to estimate the geographical variation of the association of malaria parasitaemia, diarrhoea and ARI with fever. We calculated the potential attributable fraction for each disease and country, separately, in order to estimate the percentage of potential fever cases attributed to a particular disease.

Development of acute respiratory infection led to the most increased odds of developing fever, among the childhood diseases examined, albeit diarrhoea was on average the main contributor of under-five fever across all 16 SSA countries. This dissimilarity was due to the very low prevalence of ARI compared to that of diarrhoea across all countries, which influenced the calculation of PAF. For instance in Senegal, the odds ratio of ARI was significantly higher than that of diarrhoea and yet we only estimated an approximately 7% difference in PAF between the two diseases, as the prevalence of diarrhoea was 4.5 times higher than that of ARI. Contrarily, exposure to malaria parasitaemia for under-five children was almost double to that of diarrhoea, but the corresponding odds ratio was so low, compared to the other two diseases, that the overall contribution of parasitaemia was estimated to be lower than diarrhoea and ARI. Our results supported findings of previous studies discussing the importance of ARI to fever cases (D'Acremont, 2014; Crump, 2013; Maze, 2018).

Malaria parasitaemia is not the main contributor of fever cases among under-five children in SSA. The clinical manifestation of malaria infection is individual-specific and does not necessary lead to a febrile response (Ashley, 2018). Our results were consistent with the literature as, for instance, a systematic review of the association between fever and *Plasmodium falciparum* parasitaemia in Africa found that only 22% of all fever cases were attributed to *Plasmodium* infection (D'Acremont, 2010). These finding have important implications from a public health perspective. That is, our results highlight the need for laboratory diagnosis of malaria, among children under-five with fever, and the subsequent treatment upon confirmation of the cause, which is different from the common practise of assuming malaria, when fever is present, and prescribing anti-malaria treatment.

Estimates for the contribution of the childhood disease on under-five fever based on PAF did not add up or even exceeded 100%, as in many occasions children could simultaneously experience more than one disease. For instance in Senegal, most of the under-five children experiencing ARI had also diarrhoea and thus, since only scarce cases were attributed to malaria, it was shown that most of the fever cases were both attributable to ARI and to

diarrhoea. This outcome is aligned with a recent modelling study from 21 SSA countries, which found that among under-five children with fever and a positive test for *P. falciparum* infection, only 37.7% of the fever cases were due to the malaria infection (Dalrymple, 2019). This stresses the need for further investigation of the aetiology of fever, even when it is proven that a child experiences a certain disease.

The Bayesian hierarchical models used in our study are statistical models that establish associations and do not imply causal relations, for which a causal modelling approach should have been employed. The childhood diseases data from the DHS are only available for the alive children and since the outcome of our statistical models is fever we acknowledge that this may have affected our inference.

Acknowledgments

The authors are grateful to the DHS for making the data available and all collaborating institutions whose contributions have enabled this study. The Swiss National Science Foundation (SNSF) Swiss Programme for Research on Global Issues for Development (R4D) project no. IZ01Z0-147286 funded this work. PV was funded by an ERC Advanced Grant Fellowship.

Ethical approval and consent to participate

In this study we analysed secondary data made available by the Demographic Health Survey (DHS) MEASURE. According to survey protocols and related documents of the surveys, ethical approval was obtained from the Institutional Review Board of International Consulting Firm (ICF) of Calverton, Maryland, USA, and from the national ethical committees in the countries that the surveys were contacted. Details of ethical clearance are published in the DHS reports available at https://dhsprogram.com/publications/index.cfm.

Competing interests

We declare no competing interests.

Role of funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the paper. The authors had full access to all data in the study and had final responsibility for the decision to submit for publication.

Contributors

IP conceptualised the project, processed and analysed the data, interpreted the results, and wrote the manuscript. PV acquired the financial support of the project, conceptualised the project and assisted in statistical analysis. JU and PV revised the manuscript and provided important intellectual content. All authors approved the final version of the manuscript.

Availability of Materials and Data

The study data are available upon request from the Demographic and Health Surveys program (https://dhsprogram.com/data/Access-Instructions.cfm). The “Dataset Terms of Use” does not allow us to pass on downloaded data.

Table 4.1: Descriptive analysis of the data in 16 sub-Sahara African countries, collected from the DHS programme.

	Survey year	Number of locations	Malaria parasitaemia*	Diarrhoea	ARI**	Fever	U5M***
Angola	2015	625	13.5	15.6	3.3	14.6	68
Benin	2017	540	36.3	10.4	2.8	19.4	96
Burkina Faso	2010	541	76.1	14.7	1.9	20.6	129
Burundi	2016	552	37.9	22.5	7.1	39.5	78
Cameroon	2011	577	30.0	20.9	5.4	25.9	122
Côte d'Ivoire	2011	341	41.5	17.9	3.6	23.5	108
DRC	2013	492	30.8	16.8	6.7	29.5	104
Ghana	2014	423	36.4	11.7	3.6	13.8	60
Guinea	2012	300	46.9	16.4	6.0	29.4	123
Mali	2012	413	47.2	8.6	1.6	8.6	95
Mozambique	2011	609	38.3	11.1	1.5	13.4	97
Rwanda	2014	492	7.8	12.1	5.6	18.7	50
Senegal	2016	214	0.9	15.3	3.3	12.7	51
Tanzania	2015	608	14.4	11.8	3.7	17.9	67
Togo	2013	329	37.8	15.0	3.4	21.6	88
Uganda	2016	685	30.4	19.5	9.3	33.3	64

* Measured by rapid diagnostic tests; ** Acute respiratory infection; *** under-five mortality rate (per 1,000)

Table 4.2: Posterior estimates (Odds Ratio, 95% Bayesian Credible Interval) of the association between childhood diseases and fever prevalence, adjusted for maternal, socio-economic, demographic and individual-child characteristics, as well as for health interventions. Estimates were obtained from a Bayesian geostatistical logistic regression model with spatially varying coefficients.

	Malaria parasitaemia	Diarrhoea	Acute Respiratory Infection
Angola	**2.25 (1.64, 3.08)**	**19.32 (10.32, 36.14)**	**33.35 (14.43, 76.99)**
Benin	**1.52 (1.13, 2.03)**	**9.53 (4.47, 20.29)**	**7.14 (3.37, 15.12)**
Burkina Faso	0.96 (0.71, 1.28)	**15.57 (8.47, 28.57)**	**31.42 (13.60, 72.52)**
Burundi	**3.02 (2.36, 3.87)**	**4.51 (2.62, 7.78)**	**12.56 (7.24, 21.79)**
Cameroon	**1.99 (1.51, 2.62)**	**6.75 (3.57, 12.76)**	**3.84 (2.10, 6.99)**
Côte d'Ivoire	1.01 (0.70, 1.43)	**27.48 (14.33, 52.69)**	**14.33 (6.42, 31.98)**
DRC	**1.91 (1.53, 2.40)**	**14.80 (8.06, 27.15)**	**20.37 (11.44, 36.26)**
Ghana	**1.69 (1.11, 2.56)**	**6.28 (2.69, 14.66)**	**7.73 (2.76,21.59)**
Guinea	**1.54 (1.01, 2.39)**	**14.29 (5.25, 38.47)**	**21.11 (9.02, 49.40)**
Mali	1.55 (0.93, 2.60)	**26.57 (9.20, 76.65)**	**68.15 (12.61, 367.78)**
Mozambique	**1.55 (1.16, 2.10)**	**4.17 (1.24, 13.98)**	**2.66 (7.72, 22.38)**
Rwanda	**3.48 (1.64, 7.34)**	**28.75 (13.83, 59.72)**	**15.87 (8.24, 30.53)**
Senegal	4.54 (0.36, 56.68)	**43.44 (13.60, 138.54)**	**638.87 (123.16, 3309.65)**
Tanzania	**4.88 (3.10, 7.68)**	**7.84 (3.09, 19.84)**	**9.08 (3.87, 21.29)**
Togo	1.48 (0.90, 2.42)	**18.37 (9.59, 35.17)**	**21.56 (9.99, 46.53)**
Uganda	**2.61 (2.01, 3.39)**	**13.44 (7.77, 23.23)**	**7.59 (4.45, 12.95)**

Table 4.3: Country estimates (Posterior Median, 95% Bayesian Credible Interval) of the Population Attributable Fraction (PAF) measuring the contribution of malaria, diarrhoea and acute respiratory infection on fever. Estimates were obtained from a Bayesian geostatistical logistic regression model with spatially varying coefficients.

	Malaria parasitaemia	Diarrhoea	Acute Respiratory Infection
Angola	**17.56 (10.02, 26.04)**	**71.99 (56.79, 82.65)**	**65.05 (44.61, 81.86)**
Benin	**16.44 (4.81, 29.04)**	**46.99 (26.22, 64.59)**	**32.57 (15.50, 52.78)**
Burkina Faso	-3.33 (-26.42, 16.78)	**67.11 (50.56, 79.02)**	**63.95 (41.96, 80.87)**
Burundi	**41.27 (32.38, 50.28)**	**41.95 (25.86, 57.22)**	**63.33 (48.57, 75.46)**
Cameroon	**24.47 (13.68, 34.77)**	**51.02 (33.03, 67.81)**	**28.75 (14.61, 46.83)**
Côte d'Ivoire	0.25 (-15.94, 16.13)	**81.30 (70.05, 89.81)**	**52.38 (32.13, 73.46)**
DRC	**24.36 (15.89, 32.73)**	**67.33 (50.83, 79.53)**	**68.61 (53.07, 80.05)**
Ghana	**22.88 (3.54, 39.27)**	**40.72 (18.52, 63.99)**	**33.46 (10.56, 57.90)**
Guinea	**21.09 (0.27, 38.93)**	**67.26 (39.95, 85.68)**	**70.92 (48.88, 85.08)**
Mali	20.19 (-1.85, 42.14)	**71.79 (45.19, 87.99)**	**72.40 (32.97, 94.41)**
Mozambique	**17.00 (6.02,29.48)**	**24.99 (2.26, 58.79)**	**24.61 (7.00, 51.83)**
Rwanda	**17.29 (5.35, 35.34)**	**77.72 (61.94, 87.56)**	**64.35 (47.50, 77.49)**
Senegal	6.88 (-1.44, 54.93)	**88.99 (71.79, 96.04)**	**97.16 (85.76, 99.42)**
Tanzania	**33.33 (21.71, 46.51)**	**43.97 (19.54, 68.06)**	**32.42 (14.16, 54.56)**
Togo	15.92 (-4.18, 37.50)	**72.79 (58.21, 83.97)**	**73.49 (56.58, 86.52)**
Uganda	**34.73 (25.45, 44.06)**	**70.90 (57.03, 81.28)**	**52.75 (36.68, 66.71)**

Figure 4.1: Geographical distribution of the association between malaria parasitaemia and fever prevalence, among under-five children, in 16 sub-Saharan African countries. Estimates were obtained from a Bayesian geostatistical logistic regression model with spatially varying coefficients. Estimates are presented in terms of Odds Ratio.

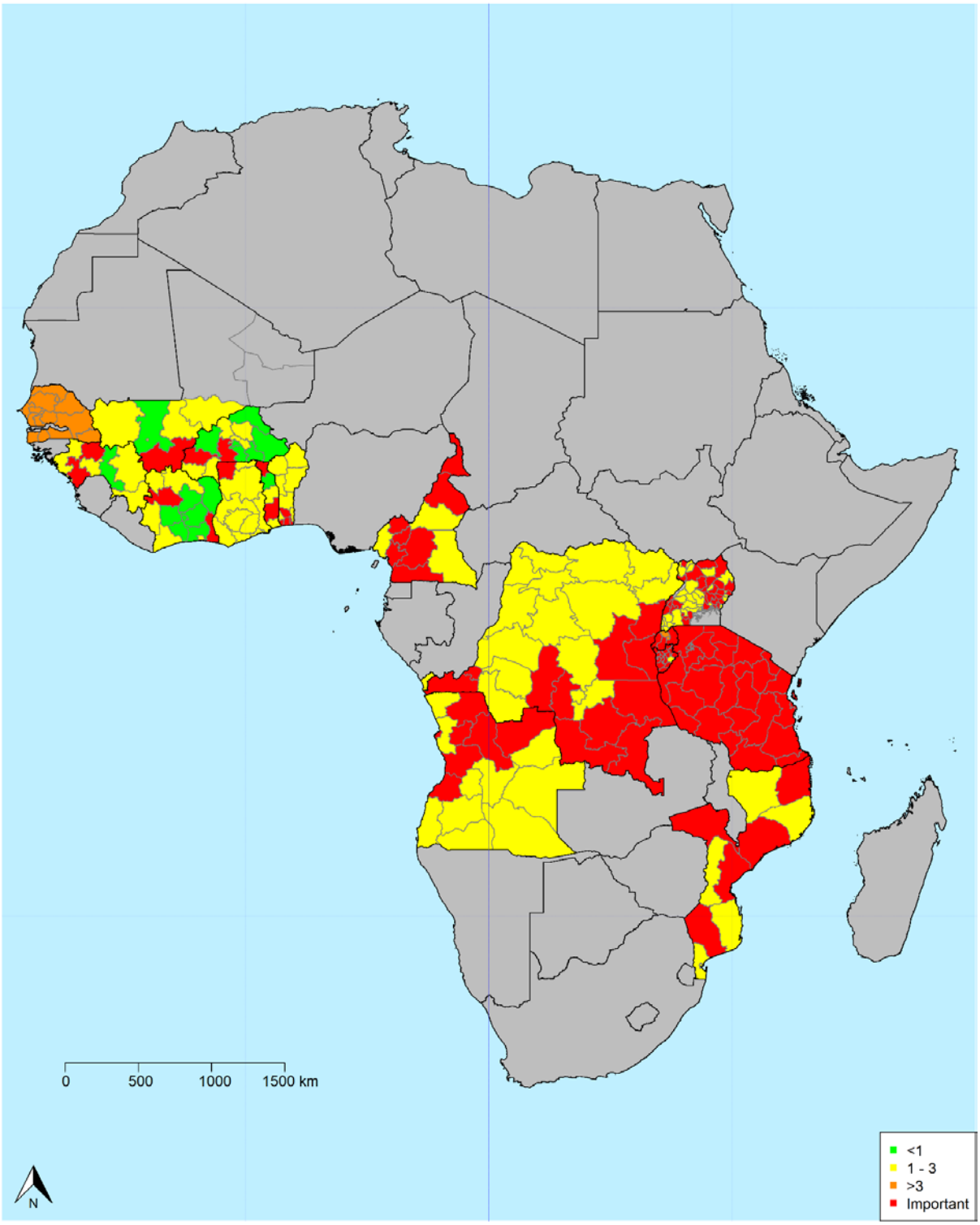


Figure 4.2: Geographical distribution of the association between diarrhoea and fever prevalence, among under-five children, in 16 sub-Saharan African countries. Estimates were obtained from a Bayesian geostatistical logistic regression model with spatially varying coefficients. Estimates are presented in terms of Odds Ratio.

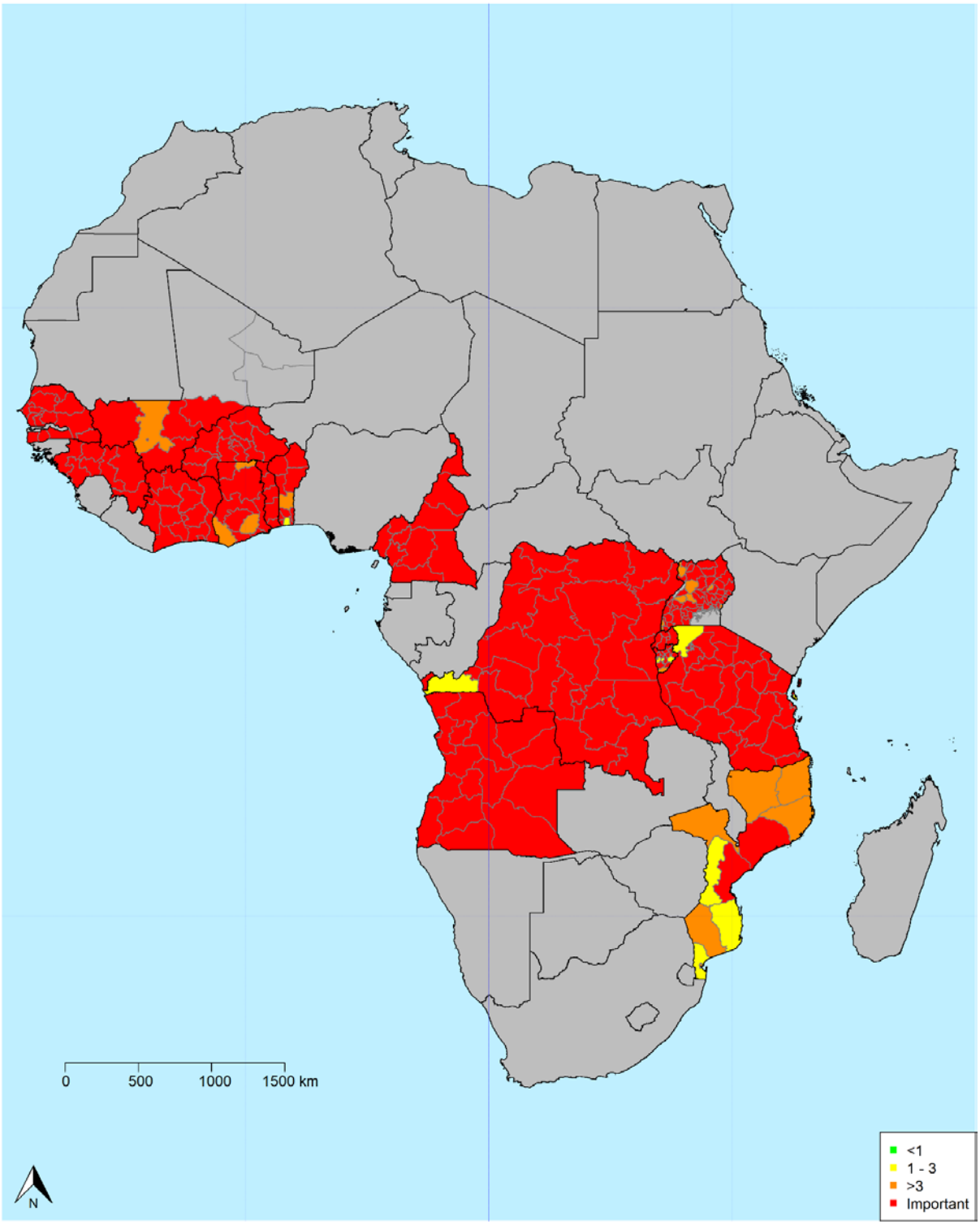
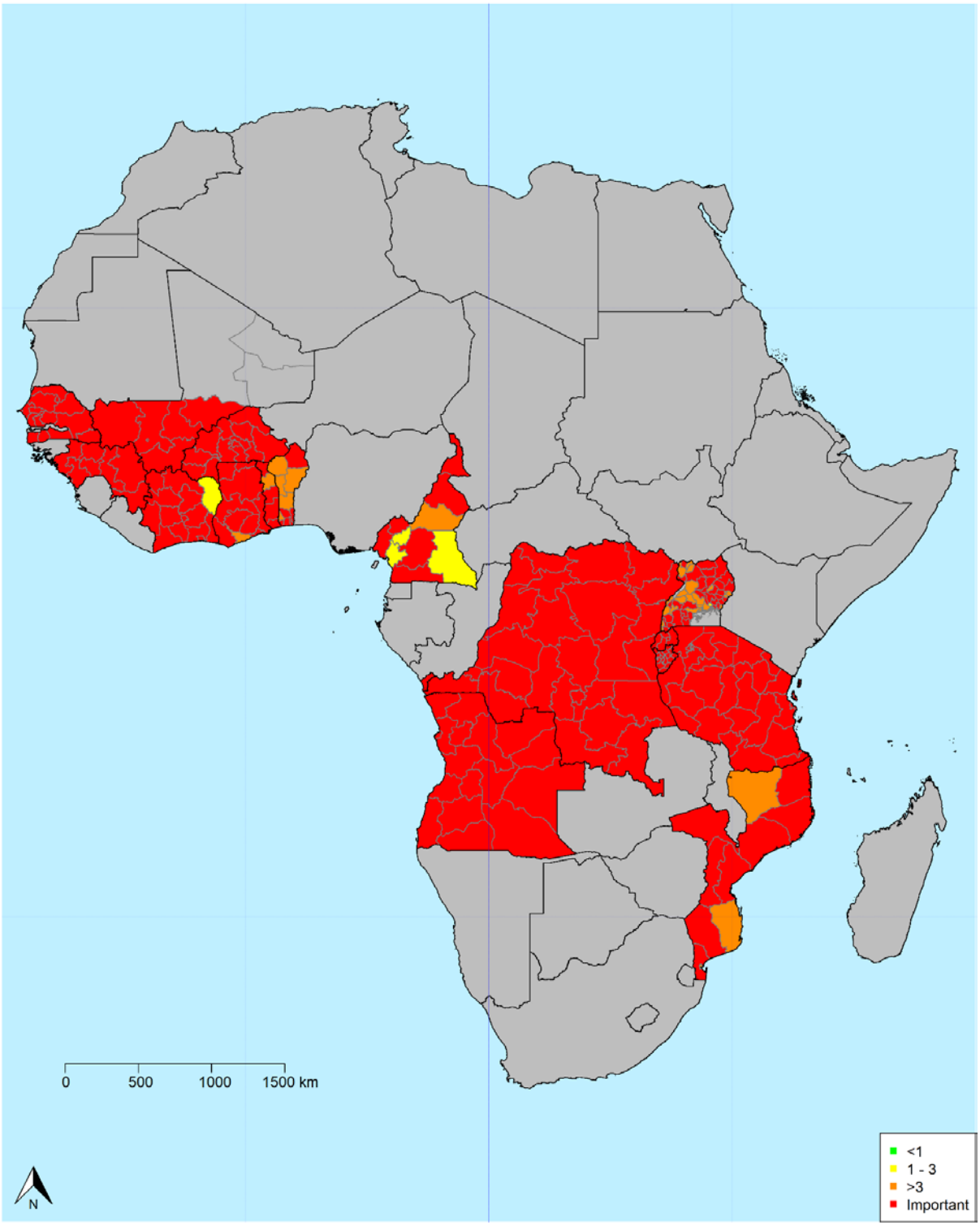


Figure 4.3: Geographical distribution of the association between acute respiratory infection and fever prevalence, among under-five children, in 16 sub-Saharan African countries. Estimates were obtained from a Bayesian geostatistical logistic regression model with spatially varying coefficients. Estimates are presented in terms of Odds Ratio.



4.5 Appendix

Model specification

The outcome of the Bayesian geostatistical logistic regression model is a binary indicator, which takes the values 0 (no fever over the last 2 weeks preceding the survey) and 1 (presence of fever). Specifically, it is represented as Y_{ji}, with j corresponding to each under-five child for a given country and i corresponding to one of the n surveyed locations of a given country, i.e. $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$. Y_{ji} follows a Bernoulli distribution with probability of success p_{ji}, i.e. the probability of developing febrile response. p_{ji} is linked to the linear predictor using the standard logit-link function as follows:

$$logit(p_{ji}) = \log\left(\frac{p_{ji}}{1-p_{ji}}\right) = \beta_0 + \sum_{\kappa=1}^{K} \beta_\kappa x_{ji\kappa} + \xi_i,$$

where $\beta = (\beta_0, \beta_1, \dots, \beta_\kappa)^T$ is the vector of K regression coefficients, $x_{ji\kappa}$ represents the value of the κ-th predictor for the j-th child at the i-th location and $\xi(s) = (\xi_1, \xi_2, \dots, \xi_n)^\tau$ is a zero-mean, multivariate Gaussian distributed random vector with Matérn covariance matrix that accounts for the spatial correlation in the response. The effect of the geostatistical intercept is cluster-specific and so each effect corresponds to each of the unique surveyed locations, i.e. $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$. The Matérn specification is achieved with the use of Matérn function as a measure that governs the covariance between any two locations s_1 and s_2. Specifically, $\xi(s) \sim N(0, \Sigma_1)$ and $\Sigma_1(s_1, s_2) = \frac{\sigma_1^2(\kappa_1\, d(s_1,s_2))^{\nu} K_\nu(\kappa_1\, d(s_1,s_2))}{\Gamma(\nu)\, 2^{\nu-1}}$, with σ_1^2 being the spatial process variance, $d(s_1, s_2)$ the distance between locations s_1 and s_2 and κ_1 the scaling parameter. The modified Bessel function of second kind and order ν is represented by K_ν in the Matérn's function formula. The Matérn specification defines the spatial range r_1 as $r_1 = \frac{\sqrt{8}}{\kappa_1}$ that particularly determines the distance at which the spatial correlation becomes negligible. In our work, we utilized an extension of the above model so that the predictor representing the disease exposure x_{ji1} is also assigned a spatially varying component, in addition to the regression coefficient β_1, which represents the national estimate of the

disease exposure-fever association. Assuming that, for a given country, there are M different provinces at which the effect of exposure would be estimated by the spatially varying effect, the model formulation would be:

$$logit(p_{ji}) = \log\left(\frac{p_{ji}}{1-p_{ji}}\right) = \beta_0 + \left(\beta_1 + \varepsilon_{1\,m(i)}\right) x_{ji1} + \sum_{\kappa=2}^{K} \beta_\kappa x_{ji\kappa} + \xi_i,$$

where now β_1 represents the national effect of the disease exposure on fever, $\varepsilon_1 = (\varepsilon_{11}, \dots, \varepsilon_{1M})^T$ are the spatially varying effects of β_1 at the different M provinces of each country and $m(i)$ is a filter of which of the M provinces the j-th child at the i-th location belongs. The vector of spatially varying effects ε_1 follows a zero-mean, multivariate Gaussian distribution with exchangeable structure:

$$\pi(\varepsilon_1|\tau) = \prod_{m=1}^{M} \frac{1}{\sqrt{2\pi}} \sqrt{\tau} \exp\left(-\frac{1}{2}\,\tau\,\varepsilon_{1m}^2\right).$$

The exchangeable structure implies that region identifiers are weighted by the values of the disease exposure and follow an exchangeable format. Under the logistic regression, the exponential function imposed on the regression coefficients provides the odds ratio. After model-fitting, a posterior sample of size 1,000 for each of the M province-specific effects was obtained. We added up the posterior sample to the national effect and summarized all estimates in terms of posterior median and 95% Bayesian credible interval. Based on these estimates we derived Figures 4.1–4.3. We also utilized the posterior sample and the province-specific effects in order to calculate the Population Attributable Function (PAF), based on the formula:

$$PAF = \frac{p\,(OR-1)}{(1\,+p\,(OR-1))},$$

where p represents the prevalence of a disease and OR represents the estimated odds ratio for a specific province. PAF can be interpreted as the percentage of all fever cases that can be attributed to a particular disease.

Additional Results

Table 4.4: Country and regional-specific estimates (Posterior Median, 95% Bayesian Credible Interval) of the Population Attributable Fraction (PAF) measuring the contribution of malaria, diarrhoea and acute respiratory infection on fever. Estimates were obtained from a Bayesian geostatistical logistic regression model with spatially varying coefficients.

Angola

	Malaria	Diarrhoea	ARI
National	17.56 (10.02, 26.04)	71.99 (56.79, 82.65)	65.05 (44.61, 81.86)
Regions			
Bengo	8.88 (-2.84, 31.00)	57.60 (19.39, 84.71)	51.43 (8.60, 91.10)
Benguela	24.03 (10.13, 42.82)	80.64 (58.15, 92.06)	87.12 (50.82, 97.30)
Bié	11.89 (-7.09, 32.44)	66.48 (32.23, 88.08)	40.71 (7.71, 81.68)
Cabinda	12.95 (-2.29, 34.21)	66.95 (34.25, 88.25)	67.76 (22.81, 92.27)
Cuando Cubango	15.12 (-7.89, 41.90)	74.71 (41.12, 91.84)	27.84 (3.47, 78.85)
Cuanza Norte	22.00 (3.79, 40.60)	70.13 (39.30, 87.79)	37.88 (6.08, 74.30)
Cuanza Sul	41.95 (20.97, 61.44)	76.50 (48.08, 90.61)	52.88 (12.65, 89.61)
Cucene	0.62 (-0.12, 2.91)	64.80 (27.51, 87.84)	53.47 (10.18, 89.98)
Huíla	3.96 (-0.76, 16.51)	77.56 (44.08, 91.99)	70.69 (25.40, 93.72)
Huambo	2.59 (-0.01, 8.75)	75.77 (47.01, 92.05)	74.64 (36.05, 94.16)
Luanda	13.12 (2.11, 32.48)	67.70 (41.31, 84.08)	66.92 (31.04, 90.22)
Lunda Norte	22.87 (3.99, 45.79)	77.27 (50.58, 92.66)	74.11 (37.27, 93.29)
Lunda Sul	5.24 (-2.47, 22.00)	64.31 (31.53, 87.89)	51.60 (9.98, 91.00)
Malanje	29.94 (15.16, 46.53)	88.60 (72.60, 95.69)	84.86 (48.54, 97.04)
Moxico	15.59 (-13.39, 43.84)	51.86 (15.97, 83.52)	55.44 (10.19, 91.62)
Namibe	0.88 (-0.37, 4.85)	73.26 (40.83, 89.36)	80.08 (39.13, 95.91)
Uíge	33.99 (14.44, 54.05)	65.13 (32.63, 86.41)	71.89 (26.89, 94.13)
Zaire	13.22 (-0.98, 34.23)	57.62 (22.88, 83.37)	44.71 (6.39, 88.78)

Benin

	Malaria	Diarrhoea	ARI
National	16.44 (4.81, 29.04)	46.99 (26.22, 64.59)	32.57 (15.50, 52.78)
Regions			
Alibori	12.89 (-8.73, 36.50)	81.86 (44.93, 95.13)	82.57 (44.75, 96.16)
Atakora	10.10 (-23.15, 37.93)	65.51 (27.05, 88.76)	21.85 (-0.93, 70.32)
Atlantique	13.12 (0.75, 28.78)	10.57 (-3.64, 42.04)	23.48 (3.04, 54.88)
Borgou	20.66 (-5.70, 44.72)	37.17 (2.86, 79.15)	17.18 (-0.88, 57.16)
Collines	5.58 (-19.07, 31.61)	21.14 (-1.82, 66.84)	17.08 (-2.82, 53.47)
Donga	14.58 (-9.00, 37.85)	53.84 (11.39, 86.59)	8.29 (-2.19, 41.70)
Kouffo	8.95 (-20.13, 37.71)	31.82 (1.80, 73.02)	23.38 (-0.42, 65.16)
Littoral	1.77 (-1.82, 9.91)	59.60 (18.15, 91.72)	32.96 (1.44, 78.57)
Mono	24.93 (3.43, 48.78)	39.21 (0.98, 80.85)	38.93 (4.30, 76.44)
Ouémé	13.01 (0.24, 29.30)	39.24 (8.16, 72.86)	23.67 (0.75, 69.96)
Plateau	13.54 (-4.27, 33.24)	55.82 (17.19, 85.06)	50.48 (8.11, 87.54)
Zou	27.04 (2.88, 48.97)	43.58 (10.63, 77.19)	52.71 (17.34, 81.58)

Burkina Faso

	Malaria	Diarrhoea	ARI
National	-3.33 (-26.42, 16.78)	67.11 (50.56, 79.02)	63.95 (41.96, 80.87)
Regions			
Boucle du Mouhoun	-14.41 (-58.49, 19.75)	76.05 (46.10, 90.95)	48.82 (6.62, 89.07)

Cascades	5.67 (-22.34, 32.28)	53.14 (17.67, 82.90)	86.22 (49.12, 97.23)
Centre-Est	-33.04 (-72.57, 6.25)	52.63 (21.17, 79.34)	38.55 (6.61, 80.38)
Centre-Nord	-30.83 (-75.41, 6.26)	69.92 (37.34, 89.35)	65.73 (16.91, 92.15)
Centre-Ouest	3.46 (-35.60, 33.40)	57.54 (23.38, 81.27)	27.79 (3.24, 77.68)
Centre-Sud	35.21 (4.38, 56.82)	78.90 (55.74, 91.11)	40.37 (6.87, 79.31)
Centre	-4.71 (-27.28, 20.57)	75.07 (46.93, 91.27)	57.04 (13.20, 88.62)
Est	-25.52 (-86.59, 21.07)	45.89 (13.20, 80.37)	53.26 (8.59, 89.28)
Haut-Bassins	36.81 (14.74, 56.33)	73.44 (45.69, 89.53)	69.02 (22.56, 92.87)
Nord	12.73 (-26.07, 38.87)	85.46 (63.03, 95.25)	72.56 (36.21, 93.37)
Plateau-Central	2.44 (-34.63, 32.20)	68.03 (31.54, 88.46)	64.82 (26.02, 89.93)
Sahel	-27.23 (-60.91, 10.03)	51.83 (13.89, 83.86)	70.33 (32.04, 91.44)
Sud-Ouest	23.43 (-15.45, 51.28)	61.95 (21.18, 86.93)	51.34 (8.99, 89.23)

Burundi

	Malaria	Diarrhoea	ARI
National	41.27 (32.38, 50.28)	41.95 (25.86, 57.22)	63.33 (48.57, 75.46)
Regions			
Bubanza	35.18 (13.19, 55.85)	50.04 (13.51, 78.38)	57.41 (22.35, 84.39)
Bujumbura Mairie	5.84 (0.64, 16.57)	45.48 (10.98, 77.22)	43.22 (11.84, 74.99)
Bujumbura Rural	12.20 (2.87, 26.67)	14.97 (-7.76, 47.84)	43.51 (10.64, 75.14)
Bururi	22.08 (8.53, 40.78)	61.71 (33.27, 80.42)	84.13 (65.60, 94.13)
Cankuzo	44.87 (22.93, 62.44)	27.13 (0.38, 63.30)	46.97 (13.27, 75.68)
Cibitoke	42.40 (21.74, 61.90)	63.28 (26.89, 86.28)	64.50 (32.18, 87.19)
Gitega	43.29 (26.79, 59.81)	47.61 (18.12, 72.20)	72.42 (47.75, 88.16)
Karuzi	47.52 (29.54, 63.03)	34.37 (3.04, 68.95)	48.70 (15.90, 80.95)
Kayanza	35.93 (15.63, 55.98)	48.12 (16.06, 75.46)	71.30 (44.65, 86.57)
Kirundo	67.73 (55.20, 76.84)	72.53 (43.87, 88.37)	82.82 (60.93, 93.40)
Makamba	33.37 (14.12, 53.35)	27.40 (-1.89, 62.47)	47.40 (12.98, 81.57)
Muramvya	34.91 (16.98, 52.72)	43.01 (7.73, 74.83)	74.18 (43.16, 90.10)
Muyinga	65.02 (51.27, 76.95)	61.32 (30.24, 83.09)	74.54 (42.27, 90.38)
Mwaro	21.60 (5.32, 41.85)	6.89 (-12.90, 42.04)	49.80 (14.15, 79.44)
Ngozi	60.27 (42.90, 72.85)	58.07 (27.15, 80.59)	66.90 (38.90, 87.24)
Rutana	41.48 (21.85, 59.09)	44.10 (8.77, 75.37)	63.34 (27.86, 86.32)
Ruyigi	10.70 (-13.75, 32.72)	5.09 (-10.50, 39.65)	28.99 (6.07, 67.13)

Cameroon

	Malaria	Diarrhoea	ARI
National	24.47 (13.68, 34.77)	51.02 (33.03, 67.81)	28.75 (14.61, 46.83)
Regions			
Adamaoua	23.32 (-0.61, 43.94)	39.70 (12.36, 69.60)	14.42 (-2.35, 53.15)
Centre	27.84 (11.29, 44.68)	53.96 (29.20, 75.83)	33.40 (10.94, 57.93)
Est	6.57 (-20.81, 32.28)	35.62 (5.30, 69.31)	8.66 (-6.57, 40.49)
Extrême-Nord	29.81 (7.89, 51.46)	71.80 (45.90, 87.57)	44.06 (15.55, 70.78)
Littoral	14.43 (2.22, 31.10)	42.68 (18.21, 70.29)	17.85 (-0.28, 48.74)
Nord-Ouest	18.24 (2.17, 39.07)	25.89 (4.40, 56.86)	23.48 (0.66, 56.75)
Nord	22.61 (3.64, 41.10)	75.86 (55.04, 87.89)	52.18 (21.98, 75.85)
Ouest	26.57 (7.40, 45.98)	35.02 (6.58, 66.87)	7.80 (-9.50, 40.89)
Sud-Ouest	54.28 (32.50, 70.22)	42.49 (14.72, 72.97)	43.65 (11.97, 77.45)
Sud	11.41 (-16.19, 35.46)	47.97 (16.28, 75.94)	32.64 (2.99, 71.73)

Côte d'Ivoire

	Malaria	Diarrhoea	ARI
National	0.25 (-15.94, 16.13)	81.30 (70.05, 89.81)	52.38 (32.13, 73.46)
Regions			
Abidjan	2.60 (-7.23, 19.63)	78.24 (55.42, 90.41)	69.39 (34.52, 88.39)
Bas-Sassandra	4.09 (-28.79, 37.62)	69.83 (38.51, 89.04)	44.70 (7.14, 82.18)
Comoé	43.69 (16.44, 62.92)	87.52 (74.41, 94.52)	74.22 (43.38, 92.24)
Denguélé	14.88 (-12.13, 48.26)	78.10 (47.92, 92.76)	37.01 (5.33, 82.27)
Gôh-Djiboua	-31.48 (-58.29, 6.54)	66.98 (30.25, 87.24)	12.13 (0.32, 53.11)
Lacs	-5.21 (-33.41, 25.31)	79.38 (53.70, 92.15)	42.51 (6.06, 86.72)
Lagunes	-13.13 (-41.51, 20.76)	78.96 (51.03, 92.84)	61.22 (16.68, 90.53)
Montagnes	16.05 (-11.98, 43.20)	80.53 (56.53, 91.57)	43.07 (11.95, 75.59)
Sassandra-Marahoué	-2.71 (-26.87, 26.73)	88.77 (73.95, 95.51)	68.41 (33.89, 88.03)
Savanes	11.69 (-14.91, 39.44)	82.89 (61.35, 93.41)	43.28 (7.22, 84.16)
Vallée du Bandama	-14.07 (-28.04, 7.48)	77.36 (47.96, 92.08)	21.53 (0.63, 68.23)
Woroba	36.56 (6.60, 60.02)	88.75 (73.41, 95.52)	78.50 (43.16, 94.49)
Yamoussoukro	-11.98 (-31.67, 25.27)	76.21 (42.48, 92.83)	27.76 (2.06, 80.27)
Zanzan	-3.37 (-34.17, 26.24)	78.00 (53.04, 92.63)	16.74 (-2.46, 60.96)

Democratic Republic of the Congo

	Malaria	Diarrhoea	ARI
National	24.36 (15.89, 32.73)	67.33 (50.83, 79.53)	68.61 (53.07, 80.05)
Regions			
Équateur	8.64 (-6.66, 25.46)	58.25 (20.21, 84.76)	63.99 (34.21, 83.32)
Bandundu	10.03 (-2.12, 24.79)	58.59 (23.28, 84.51)	54.42 (24.37, 79.89)
Bas-Congo	31.67 (9.54, 50.04)	26.01 (-1.84, 62.62)	48.15 (15.84, 79.05)
Kasaï-Occidental	39.38 (16.09, 57.49)	76.18 (45.79, 91.85)	60.64 (29.79, 82.93)
Kasaï-Oriental	20.56 (-2.23, 40.56)	88.33 (73.20, 95.13)	79.59 (59.36, 90.82)
Katanga	36.01 (15.04, 52.76)	76.16 (47.99, 92.17)	62.05 (33.78, 83.07)
Kinshasa City	30.90 (13.40, 49.13)	67.78 (24.00, 89.47)	45.96 (16.24, 77.47)
Kivu	23.48 (8.59, 39.72)	78.61 (53.69, 90.73)	83.12 (66.83, 91.60)
Orientale	11.84 (-9.34, 31.52)	56.75 (22.09, 82.86)	66.67 (39.04, 85.55)

Ghana

	Malaria	Diarrhoea	ARI
National	22.88 (3.54, 39.27)	40.72 (18.52, 63.99)	33.46 (10.56, 57.90)
Regions			
Ashanti	18.53 (-2.59, 44.10)	40.67 (6.38, 74.58)	34.34 (3.49, 75.56)
Brong Ahafo	24.27 (-5.80, 52.27)	44.42 (7.43, 76.79)	35.58 (5.55, 76.66)
Central	26.74 (-9.93, 53.37)	34.42 (3.61, 71.68)	20.75 (-1.53, 62.79)
Eastern	9.23 (-20.69, 39.72)	26.79 (-0.02, 62.88)	35.06 (4.32, 70.77)
Greater Accra	11.40 (-1.61, 31.94)	35.36 (2.69, 74.60)	31.66 (0.8, 74.89)
Northern	1.00 (-37.33, 36.33)	63.02 (23.11, 88.31)	35.01 (4.14, 75.76)
Upper East	15.76 (-3.84, 39.52)	29.71 (-0.01, 71.41)	29.32 (0.46, 72.17)
Upper West	45.29 (9.44, 69.12)	51.71 (12.38, 82.83)	20.72 (1.07, 66.67)
Volta	24.61 (-4.08,51.25)	43.11 (8.98, 82.14)	44.63 (5.30, 84.67)
Western	25.99 (-10.97,55.56)	20.71 (-0.01, 61.59)	48.88 (9.70, 84.44)

Guinea

	Malaria	Diarrhoea	ARI
National	21.09 (0.27, 38.93)	67.26 (39.95, 85.68)	70.92 (48.88, 85.08)
Regions			
Boké	22.23 (-7.37, 49.29)	55.08 (9.56, 87.87)	66.94 (25.88, 89.03)
Conakry	1.42 (-1.86, 11.33)	72.74 (35.82, 92.15)	33.20 (2.22, 75.51)
Faranah	-15.25 (-55.20, 26.16)	61.16 (20.59, 87.41)	56.75 (13.46, 87.11)
Kankan	26.49 (-4.85, 51.29)	81.21 (43.34, 95.33)	87.85 (57.60, 97.35)
Kindia	28.31 (1.32, 53.49)	82.17 (47.72, 95.84)	85.74 (59.71, 95.49)
Labé	28.71 (1.70, 54.82)	82.14 (41.89, 94.85)	85.33 (62.50, 94.99)
Mamou	23.82 (-5.44, 54.15)	40.90 (0.66, 78.96)	50.17 (9.10, 83.36)
Nzérékoré	11.41 (26.43, 44.57)	45.22 (4.66, 83.43)	71.26 (22.80, 94.57)

Mali

	Malaria	Diarrhoea	ARI
National	20.19 (-1.85, 42.14)	71.79 (45.19, 87.99)	72.40 (32.97, 94.41)
Regions			
Bamako	3.10 (-2.62, 17.65)	43.26 (1.00, 84.13)	74.45 (20.87, 96.24)
Kayes	19.19 (-6.81, 48.23)	70.69 (18.16, 96.05)	70.42 (16.76, 96.20)
Koulikoro	-5.34 (-34.73, 25.68)	30.71 (-0.01, 72.87)	59.90 (11.10, 93.67)
Mopti	30.93 (-12.20, 59.01)	84.27 (48.19, 96.65)	76.65 (24.71, 96.33)
Ségou	21.89 (-13.16, 51.92)	80.24 (41.91, 95.32)	73.38 (22.07, 96.26)
Sikasso	40.34 (7.28, 63.29)	90.37 (60.52, 98.45)	76.36 (25.38, 96.39)
Kidal	–	–	–
Tombouctou	–	–	–
Gao	–	–	–

Mozambique

	Malaria	Diarrhoea	ARI
National	17.00 (6.02,29.48)	24.99 (2.26, 58.79)	24.61 (7.00, 51.83)
Regions			
Cabo Delgado	28.15 (4.67, 48.08)	11.18 (-2.39, 54.37)	18.24 (0.56, 66.05)
Gaza	20.37 (1.76, 40.43)	33.37 (-0.23, 78.31)	32.43 (1.49, 78.06)
Inhambane	4.96 (-15.94, 27.05)	5.81 (2.38, 44.28)	12.35 (-0.91, 55.42)
Manica	15.06 (-4.82, 37.03)	11.72 (-5.46, 61.86)	25.18 (1.58, 68.76)
Maputo City	0.78 (-0.55, 3.83)	6.67 (-7.53, 48.26)	30.80 (4.87, 68.41)
Maputo	3.50 (-2.96, 16.13)	20.27 (-5.41, 69.62)	36.56 (0, 76.38)
Nampula	4.32 (-18.50, 27.88)	23.03 (-3.54, 69.75)	28.16 (1.31, 77.50)
Nassa	15.52 (-8.36, 38.42)	24.03 (-4.26, 72.65)	12.30 (-1.16, 53.26)
Sofala	32.20 (15.48, 49.75)	59.43 (17.03, 86.64)	35.44 (6.46, 76.21)
Tete	21.08 (0.68, 42.01)	26.41 (-3.34, 69.50)	18.18 (1.05, 60.49)
Zambezia	23.55 (0.17, 45.71)	59.47 (20.42, 86.97)	18.20 (1.10, 62.39)

Rwanda

	Malaria	Diarrhoea	ARI
National	17.29 (5.35, 35.34)	77.72 (61.94, 87.56)	64.35 (47.50, 77.49)
Regions			
Amajyaruguru	2.35 (0, 15.44)	74.81 (46.96, 91.09)	70.62 (36.77, 89.74)
Amajyepfo	20.03 (7.49, 36.24)	76.26 (54.79, 89.45)	52.37 (30.47, 72.47)
Iburasirazuba	32.47 (14.32, 52.27)	84.51 (66.71, 93.25)	72.29 (42.94, 88.88)
Iburengerazuba	8.04 (0.69, 25.18)	78.77 (56.83, 90.42)	76.65 (53.62, 90.15)
Umujyi wa Kigali	2.68 (-0.55, 16.01)	66.89 (29.96, 88.54)	43.41 (12.57, 71.48)

Senegal

	Malaria	Diarrhoea	ARI
National	6.88 (-1.44, 54.93)	88.99 (71.79, 96.04)	97.16 (85.76, 99.42)
Regions			
Dakar	3.33 (-0.83, 53.84)	89.53 (61.64, 97.77)	98.84 (90.01, 99.78)
Diourbel	0.57 (-0.14, 15.56)	80.19 (36.11, 96.22)	92.13 (53.94, 99.09)
Fatick	0.68 (-0.18, 16.50)	80.33 (39.55, 95.33)	96.66 (72.98, 99.60)
Kédougou	27.30 (-9.39, 85.19)	83.04 (46.37, 96.94)	96.76 (75.97, 99.59)
Kaffrine	2.24 (-0.80, 44.55)	95.13 (80.84, 98.87)	98.69 (89.93, 99.81)
Kaolack	4.28 (-1.00, 58.11)	96.26 (84.38, 99.17)	98.87 (88.76, 99.81)
Kolda	29.35 (-3.87, 88.32)	87.99 (62.00, 96.82)	97.91 (84.30, 99.76)
Louga	3.35 (-0.83, 51.68)	88.92 (46.68, 98.51)	96.90 (74.25, 99.64)
Matam	4.01 (-1.15, 58.33)	83.56 (41.50, 96.81)	94.60 (59.27, 99.38)
Sédhiou	0.83 (-0.02, 19.49)	88.49 (57.17, 97.52)	97.12 (79.19, 99.60)
Saint–Louis	3.24 (-0.84, 52.79)	83.97 (39.11, 97.62)	97.20 (81.30, 99.72)
Tambacounda	9.61 (-2.46, 73.63)	88.47 (57.34, 97.96)	96.15 (75.55, 99.57)
Thiès	1.24 (-0.25, 27.02)	74.98 (24.14, 95.82)	96.85 (76.29, 99.62)
Ziguinchor	1.82 (-0.55, 38.22)	92.02 (65.90, 98.62)	98.23 (86.11, 99.80)

Tanzania

	Malaria	Diarrhoea	ARI
National	33.33 (21.71, 46.51)	43.97 (19.54, 68.06)	32.42 (14.16, 54.56)
Regions			
Arusha	3.81 (0.53, 11.87)	38.01 (5.41, 78.15)	30.27 (3.18, 69.67)
Dar es Salaam	4.81 (0.72, 15.58)	44.90 (5.02, 79.35)	52.40 (15.24, 83.45)
Dodoma	3.59 (0.50, 13.09)	45.27 (4.80, 81.30)	23.25 (2.27, 65.75)
Geita	50.26 (27.65, 69.75)	28.80 (1.95, 71.50)	19.98 (1.20, 57.16)
Iringa	3.84 (0.58, 13.12)	30.50 (1.67, 74.90)	40.79 (7.06, 78.98)
Kagera	49.17 (23.40, 68.58)	23.87 (-1.84, 63.70)	34.35 (6.68, 70.33)
Katavi	32.88 (10.26, 58.48)	42.24 (5.33, 79.21)	15.18 (1.70, 49.97)
Kigoma	56.91 (34.95, 74.04)	59.58 (17.39, 86.97)	48.02 (10.37, 82.02)
Kilimanjaro	3.60 (0.65, 12.76)	47.42 (10.32, 83.92)	53.59 (15.68, 84.87)
Lindi	44.90 (16.70, 71.39)	49.38 (10.06, 82.03)	24.49 (2.50, 62.21)
Manyara	3.49 (0.56, 12.49)	35.14 (1.10, 75.75)	11.10 (1.10, 44.53)
Mara	51.44 (26.44, 72.32)	63.52 (20.98, 89.11)	31.34 (3.26, 71.29)
Mbeya	3.16 (0.45, 11.43)	46.32 (4.82, 83.46)	34.71 (4.64, 75.91)
Morogoro	46.54 (17.89, 70.90)	39.67 (4.40, 80.01)	38.55 (6.53, 78.60)
Mtwara	54.53 (24.83, 75.90)	49.54 (4.85, 83.18)	46.18 (10.10, 82.58)
Mwanza	36.57 (12.87, 62.31)	44.36 (12.09, 79.24)	34.74 (5.71, 69.60)
Njombe	2.48 (0.36, 9.30)	38.85 (4.00, 79.56)	41.90 (8.00, 80.41)
Pemba North	3.73 (0.58, 11.89)	47.14 (7.94, 83.32)	39.84 (6.29, 79.33)
Pemba South	3.70 (0.46, 12.99)	56.28 (14.50, 87.07)	33.57 (5.77, 74.96)

Pwani	20.89 (0.34, 48.00)	42.08 (6.35, 80.45)	29.85 (2.55, 71.94)
Rukwa	9.88 (1.73, 28.30)	62.40 (19.81, 88.40)	35.54 (6.50, 75.04)
Ruvuma	46.95 (19.46, 70.74)	55.05 (14.71, 86.26)	4.02 (0.39, 22.29)
Shinyanga	48.19 (23.17, 71.48)	43.74 (8.38, 79.78)	30.80 (5.74, 72.90)
Simiyu	40.89 (15.28, 65.91)	52.69 (13.62, 83.76)	34.80 (7.87, 74.64)
Singida	27.99 (8.68, 55.50)	37.84 (2.95, 79.66)	31.01 (4.30, 70.75)
Tabora	39.49 (12.74, 65.21)	26.82 (3.14, 67.72)	15.18 (1.49, 48.14)
Tanga	13.76 (2.58, 37.29)	29.14 (2.55, 70.88)	26.59 (3.67, 67.40)
Zanzibar North	3.76 (0.60, 12.40)	31.64 (1.58, 74.68)	31.07 (2.81, 72.28)
Zanzibar South and Central	1.94 (0.28, 6.56)	25.53 (-0.18, 67.84)	14.72 (0.71, 53.78)
Zanzibar West	3.69 (0.57, 11.27)	43.12 (5.88, 80.41)	31.99 (5.11, 72.03)

Togo

	Malaria	Diarrhoea	ARI
National	15.92 (-4.18, 37.50)	72.79 (58.21, 83.97)	73.49 (56.58, 86.52)
Regions			
Centre	16.29 (-19.85, 47.97)	73.69 (38.90, 92.50)	69.14 (27.38, 92.85)
Kara	-8.02 (-44.65, 32.84)	65.05 (32.24, 85.84)	25.16 (-0.32, 73.51)
Maritime	7.99 (-6.65, 27.62)	28.39 (5.38, 58.77)	66.12 (39.45, 84.79)
Plateaux	30.87 (1.07, 55.52)	58.31 (25.48, 80.62)	67.66 (37.08, 87.42)
Savanes	33.18 (2.78, 60.18)	97.48 (93.95, 98.92)	96.67 (90.97, 98.88)

Uganda

	Malaria	Diarrhoea	ARI
National	34.73 (25.45, 44.06)	70.90 (57.03, 81.28)	52.75 (36.68, 66.71)
Regions			
Adjumani	7.12 (-2.50, 31.98)	47.42 (5.73, 89.71)	19.15 (-0.56, 66.89)
Iganga	53.90 (22.72, 72.78)	89.78 (64.81, 97.40)	74.40 (36.87, 91.27)
Jinja	1.60 (-7.83, 20.93)	35.26 (-0.60, 76.74)	26.73 (-2.50, 69.65)
Kabale	56.50 (21.73, 80.61)	84.04 (47.13, 96.02)	46.68 (5.68, 82.89)
Kabarole	53.09 (14.91, 77.92)	85.60 (50.75, 96.80)	36.50 (-11.68, 80.81)
Kaberamaido	12.70 (-2.24, 38.48)	84.96 (46.07, 97.00)	41.44 (1.00, 82.16)
Kalangala	17.70 (-30.05, 53.75)	38.00 (-0.54, 84.58)	27.91 (-2.04, 74.04)
Kampala	63.40 (39.79, 78.29)	83.15 (50.75, 96.36)	47.29 (6.76, 86.37)
Kamuli	18.54 (-7.09, 51.32)	55.45 (7.49, 90.96)	26.99 (-1.08, 79.02)
Kamwenge	50.29 (21.77, 73.39)	74.75 (32.14, 93.58)	72.56 (28.62, 93.48)
Kanungu	30.20 (-5.93, 60.82)	59.67 (10.51, 89.09)	20.02 (-8.53, 70.94)
Apac	1.59 (-0.28, 7.95)	49.81 (4.88, 84.07)	72.25 (28.81, 94.17)
Kapchorwa	22.03 (3.09, 51.09)	90.65 (63.84, 98.43)	73.88 (29.72, 93.54)
Kasese	10.53 (-5.97, 40.23)	49.92 (-2.43, 88.38)	34.05 (-10.88, 79.90)
Katakwi	4.28 (-0.98, 20.54)	54.89 (4.18, 91.42)	23.74 (-1.99, 74.65)
Kayunga	1.00 (-0.13, 4.91)	45.44 (9.27, 78.66)	27.79 (-0.58, 69.09)
Kibale	78.20 (58.42, 88.82)	92.99 (72.24, 98.61)	86.62 (50.43, 97.14)
Kiboga	21.06 (0.19, 52.90)	72.75 (26.07, 93.73)	39.75 (-8.67, 82.29)
Kisoro	19.76 (-5.10, 57.33)	48.61 (-1.04, 90.13)	30.49 (-5.11, 80.41)
Kitgum	8.71 (-3.49, 35.92)	58.01 (3.29, 91.45)	49.17 (-2.16, 88.22)
Kotido	35.14 (10.40, 62.35)	67.33 (20.04, 90.83)	39.74 (-1.09, 79.49)
Kumi	70.24 (43.22, 84.98)	88.67 (58.58, 97.47)	82.73 (37.39, 96.33)
Arua	52.24 (0.07, 79.40)	73.54 (14.97, 94.60)	49.40 (1.70, 88.55)
Kyenjojo	14.28 (-13.22, 45.93)	38.09 (-0.84, 83.41)	6.16 (-0.17, 38.00)
Lira	14.62 (-4.94, 48.17)	42.47 (-0.78, 88.76)	39.82 (-2.87, 83.91)
Luwero	1.67 (-0.27, 8.09)	36.31 (-3.05, 82.75)	24.49 (-1.76, 75.29)

Masaka	56.94 (19.40, 77.43)	71.58 (23.97, 93.39)	71.67 (24.60, 92.27)
Masindi	43.56 (2.56, 69.04)	73.51 (29.08, 93.75)	71.85 (25.67, 93.13)
Mayuge	52.08 (25.02, 71.88)	94.81 (83.84, 98.61)	88.18 (62.65, 96.79)
Mbale	48.59 (21.04, 71.33)	64.35 (16.46, 91.14)	56.40 (15.49, 86.40)
Mbarara	–	–	–
Bugiri	–	–	–
Moroto	37.15 (8.69, 62.02)	80.04 (44.97, 94.46)	51.10 (14.28, 80.92)
Moyo	15.75 (-8.66, 49.27)	64.45 (15.18, 93.46)	32.25 (-1.88, 82.50)
Mpigi	54.29 (28.22, 77.05)	90.32 (64.30, 97.95)	83.55 (55.39, 95.40)
Mubende	-20.33 (-47.57, 17.11)	26.82 (-4.40, 77.19)	12.18 (-0.53, 60.36)
Mukono	57.83 (38.81, 73.35)	85.86 (63.18, 95.62)	84.89 (55.95, 96.22)
Nakapiripirit	15.35 (-2.44, 41.46)	59.70 (19.36, 89.48)	35.40 (3.04, 79.05)
Nakasongola	15.14 (0.45, 41.06)	48.91 (12.93, 82.31)	43.12 (6.68, 81.71)
Nebbi	32.89 (-7.47, 62.79)	54.57 (7.90, 88.55)	58.23 (11.30, 87.24)
Ntungamo	13.33 (-2.13, 45.24)	77.89 (35.15, 95.91)	22.52 (0.50, 76.38)
Pader	12.21 (-1.34, 39.86)	59.25 (13.41, 92.58)	46.67 (1.65, 86.86)
Bundibugyo	21.76 (0.15, 52.12)	77.61 (35.71, 95.98)	65.22 (21.89, 91.97)
Pallisa	34.31 (12.39, 58.29)	66.30 (30.42, 88.71)	31.27 (0.51, 70.49)
Rakai	62.81 (31.63, 83.28)	81.57 (43.60, 96.49)	67.94 (26.27, 90.98)
Rukungiri	10.73 (-1.14, 39.39)	71.39 (18.93, 94.65)	43.10 (2.69, 88.04)
Sembabule	47.01 (14.87, 73.04))	84.97 (46.86, 96.99)	75.71 (36.72, 93.76)
Sironko	9.66 (-0.76, 36.40)	80.20 (47.95, 95.23)	77.63 (45.31, 93.65)
Soroti	21.99 (-12.93, 54.93)	60.75 (15.27, 90.86)	43.39 (7.21, 83.84)
Tororo	31.64 (11.38, 56.39)	76.70 (43.52, 94.47)	48.06 (11.42, 84.52)
Wakiso	61.40 (35.44, 80.63)	84.52 (56.46, 96.29)	83.85 (56.49, 95.89)
Yumbe	18.69 (1.01, 54.70)	52.78 (8.80, 88.26)	49.87 (9.76, 87.12)
Bushenyi	18.26 (-0.63, 52.44)	63.54 (16.61, 92.75)	60.61 (13.20, 91.06)
Busia	8.56 (-2.33, 31.31)	53.00 (10.75, 89.77)	34.47 (2.34, 79.72)
Gulu	48.20 (27.08, 71.42)	86.48 (65.52, 96.04)	41.21 (4.79, 81.63)
Hoima	13.16 (1.95, 36.93)	71.33 (38.56, 92.47)	33.88 (6.28, 77.26)

Chapter 5

Effects of health interventions on changes in under-five mortality risk across sub-Saharan Africa: a geostatistical analysis

Isidoros Papaioannou[1,2], Jürg Utzinger[1,2], Penelope Vounatsou[1,2]

Affiliations:

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

Abstract

An unprecedented reduction of the under-five mortality rate (U5MR) between 1990 and 2018 was achieved in the sub-Sahara African, as pre-school mortality fell from 182 deaths per thousand live births to 78 deaths. Most of this reduction is attributed to the scaling-up of preventative and curative health interventions against some of the leading causes of under-mortality, which were implemented particularly towards the end of the Millennium Development Goals era (MDGs). It is of great public health importance to understand which of these health interventions were associated with the changes in mortality risk that drove the reductions in the U5MR. We compiled data on under-five mortality, coverage of health interventions and of mortality-related risk factors from the Demographic and Health Surveys (DHS) for 21 sub-Sahara African countries that had data for two different time points, i.e. surveys. We employed a Bayesian geostatistical Weibull survival modelling approach in order to associate health interventions with changes in mortality risk for the two time points. Our results highlighted that Bacillus Calmette–Guérin (BCG) immunization, vitamin A supplementation and deworming medication were the health interventions mostly associated with the changes in all-cause under-five mortality in sub-Saharan Africa. Our results reflected the crucial role of vaccinations, micronutrient supplementation and treatments of diseases in the reduction of under-five mortality risk in the African setting.

5.1 Introduction

Between 1990 and 2018, the world has experienced a remarkable reduction in the number of deaths for children below the age of five. The number of under-five deaths declined from 12.5 million to 5.3 while the under-five mortality rate (U5MR), i.e. the probability for a child to die before its fifth birthday, fell from 93 deaths per thousand live births to 39 deaths (UN IGME, 2019). Most of these reductions took place between 2000 and 2015, hence during the Millennium Development Goals (MDGs) era. It has been estimated that the millennium commitment led to more than 50% reduction of the U5MR. Similar reductions were observed for the countries of sub-Saharan Africa (SSA), as the U5MR fell from 182 deaths per thousand live births in 1990 to 78 in 2018. This translates into an annual reduction rate of 3% over these 28 years, albeit the highest annual rate was observed between 2000 and 2018 at 3.8%. Yet, the African region remains the epicentre of the most under-five deaths worldwide and currently oversees pre-school mortality trends that the developed countries experienced two decades ago. Crucially, the SSA is home to five out of six countries with an U5MR that exceeds 100 deaths per thousand live births (WHO, 2019c).

The scaling-up of preventative or curative health interventions mainly drove the reductions in the total number of deaths for under-five children during, and dynamically towards, the end of the MDGs period (Snow, 2015). Examples of changes in the distribution and usage of health interventions include the 30% increase in the ownership of Insecticide-Treated Nets (WHO, 2017a) (ITNs) and the changes in diphtheria-tetanus-pertussis-3 (DPT-3) vaccination from below 50% in the Africa region (Mosser, 2019) to 72.3% of second administrative unit[6]. These differences have substantially added to the protection of under-five children in SSA. Between 2000 and 2015, malaria control tools, championed by the use of ITNs to minimize the human-mosquito interaction, have been estimated to avert 663 million clinical malaria cases, by a Bayesian geostatistical model utilizing data from approximately 27000 georeferenced clusters over a 19-year period until 2014 (Bhatt, 2015). The strong impact of vaccinations, improved hygiene facilities and oral rehydration solutions to the decline of

some of the leading causes of under-five mortality has been reported by international organisations such as the United Nations (UN IGME, 2019). A recent systematic review and meta-analysis indicated a 39% reduction of neonatal mortality risk in SSA with the presence of antenatal care visit that is conducted by a skilled provider (Tekelab, 2019). Despite the proven association between health interventions and under-five mortality in the African setting, there is a scarcity of studies taking a holistic approach and associating a large set of well proven for their effectiveness health interventions with under-five mortality, in order to take into account the synergistic effect of the scaling-up of multiple interventions. Additionally, to our knowledge, there are no studies yet that they have explicitly modelled impact of interventions, in terms of statistical associations, on the changes in mortality risk in Africa. An example of such a study has been conducted for the effect of health interventions on the changes in malaria parasitaemia risk for six SSA countries and reported the protective effect of ITNs against malaria risk (Giardina, 2014).

Our geostatistical modelling study aims to associate well-proven, preventive or curative health interventions with changes in under-five mortality risk, between two time points, in 21 SSA countries. To achieve that, we compiled and analysed health-related data coming from the Demographic and Health surveys (DHS), as well as climatic and environmental data from other open-access sources. We included only the countries with at least two DHS surveys and developed Bayesian geostatistical Weibull survival models to estimate the effect of health interventions on changes in under-five mortality risk across sub-Saharan Africa. We also implemented rigorous Bayesian variable selection procedures to select the best set of interventions for subsequent model fit.

5.2 Methods

Data sources

The current study is primarily based on data coming from the cross-sectional, nationally and sub-nationally representative, household based DHS surveys. We compiled data from 21 SSA countries that had at least two DHS surveys in order to estimate effect of interventions

on the changes in mortality over that two time points. Each survey consisted of a unique dataset that contained individual-level information for children less than the age of five. The key indicators incorporated in our study were as follows: the current or death age of each child, maternal characteristics (educational attainment of the mother of the child and her age at first birth), individual child characteristics (sex, delivered by caesarean section, birth order of the child), the socio-economic status of the family at which the child was a member and a classification of the type of region that the child was residing (rural or urban). The individual-level morality data were geolocated, i.e. for each child we had available information for its residential coordinates. The set of curative or preventive health intervention available in the surveys included a wide range of health solutions, i.e. micronutrient supplementation, vaccinations, treatment of diseases, breastfeeding, malaria bednets, reproductive health and Water, Sanitation and Hygiene (WASH) practices. A full list of all interventions considered in our study is presented in Table 6.2. Importantly, the health interventions were modelled at cluster-level, since interventions are not reported by the DHS for the dead children therefore individual-level modelling for these indicators might have resulted on biased estimates.

The DHS Program maintains strict standards for protecting the privacy of respondents and household members in all DHS surveys. Before each interview or biomarker test is conducted, an informed consent statement is read to the respondent, who may accept or decline to participate. In addition, verbal informed consent for each parasitaemia test is provided by the child's parent/guardian/caregiver on behalf of children less than 5 years before the test is conducted. The interviewer requests verbal consent by reading a prescribed statement to the respondent and records in the questionnaire the reply from the respondent. The interviewer signs his or her name attesting to the fact that he/she read the consent statement to the respondent. Written consent is not included (https://dhsprogram.com/publications/publication-dhsm7-dhs-questionnaires-and-manuals.cfm).

In addition to the DHS surveys, we downloaded from open-access sources climatic and environmental data. The open-access sources from which we extracted our data were the

Global Rural and Urban Mapping Project (GRUMP), the Shuttle Radar Topography Mission (SRTM), the U.S. Geological Survey-Earth Resources Observation Systems (USGSS) Data portal and the Moderate Resolution Imaging Spectroradiometer (MODIS). We had available information on rainfall (USGSS), distance to the nearest water bodies (DWATER) (MODIS), altitude (SRTM), the normalized difference vegetation index (NDVI) (MODIS), the land surface temperature at night (LSTN) (MODIS) and the land surface temperature at day (LSTN) (MODIS) at cluster-level. We also had data on the land cover type of each cluster, i.e. cropland, grassland or forest (MODIS). Rural or urban classification for the residence of the children was based on the GRUMP project. A list of environmental and climatic data used in the study, together with their spatial resolution, is provided in the Appendix.

Statistical analysis

The core tool of our analysis is a specific type of spatial, mixed-effects model, namely the geostatistical model (Diggle, 1998). This approach assumes that the outcome of the model, in our case the under-five mortality, is correlated in space and the random intercept (geostatistical part) explains this feature of the data. The spatial random intercept was incorporated in the linear predictor as a cluster level Gaussian process with Matérn covariance matrix (Banerjee, 2014). Since the outcome of the model was the current or death age of each child, together with the corresponding censoring status, we assumed a Weibull lifetime distribution for the outcome and hence we resulted on a Weibull survival, geostatistical model. We employed Bayesian inference and hence we summarized regression coefficients by their posterior distribution. Assuming that the 95% Bayesian credible interval did not include zero, we interpreted and called the corresponding predictor as (statistically) important, which is equivalent to statistical significance under a frequentist approach. Model specification details can be found in the Appendix.

We performed a bivariate geostatistical variable selection procedure by fitting for each country and health intervention, separately, a Bayesian geostatistical Weibull survival model in order to select interventions for subsequent model fit. The procedure involved only the

most recent DHS surveys for each country, as we did not use interventions data from the first-in-time surveys. For the set of selected variables, we used the variance inflation factor in order to exclude highly correlated predictors (O'Brien, 2007). In order to estimate the effects of health interventions on changes in under-five risk we followed a 3-step process. Initially, we fitted a Weibull survival geostatistical model for the first DHS survey of each country, using the climatic and environmental data as predictors. As a second step, we extracted the climatic and environmental data for the first survey year at the locations of the second survey and subsequently used these data to predict the outcome at the locations of the second survey. Under the assumption of a Weibull survival model, the linear predictor containing the fixed and random effects were linked to the scale parameter of the Weibull lifetime distribution and hence what we predicted in the second step is essentially the scale parameter at the locations of the second survey. The final steps involved modelling of the difference in scale parameters of the two Weibull survival distributions corresponding to the two time points, by fitting the geostatistical model for the second survey using the predicted scale from the first survey as an offset. The linear predictor of the last step incorporated the difference in the climatic data between the two time points, the selected health interventions at the second time point and a set of maternal, household and individual child covariates used as adjustments. The posterior coefficients of the health interventions were summarized in order to identify associations between interventions and changes in under-five mortality risk. A detailed description of the procedures and model formulation is provided in the Appendix.

The analysis was performed in an integrated development environment of the R software (R Core Team (2019), 2019), i.e. the Rstudio (RStudio Team (2018), 2018). The bivariate geostatistical variable selection and the first two steps of the subsequent model fitting, i.e. fitting data from the first survey and predicting at locations of the second using the corresponding climatic and environmental data, were made using the Integrated Nested Laplace Approximation (Rue, 2009; Lindgren, 2011) (INLA) in order to conduct fast and reliable Bayesian Inference. The final step of the 3-process fitting was performed using the

STAN probabilistic programming language (Carpenter, 2017) and its interface with R called RStan (Stan Development Team, 2018).

5.3 Results

The DHS data compiled for this geostatistical analysis included cross-sectional information on 437,500 under-five children in 2 different time points and across 21 SSA countries. Almost equally split, 204,602 of these observations correspond to the surveys conducted at the first time point for each country. Accounting for all surveys, a total number of 20,172 unique locations represented the clusters for which we had available information on pre-school children. Nigeria, followed by Malawi, were the countries with the highest number of unique locations and observations in both sets of surveys, i.e. DHS that took place chronologically first and second. Benin and Sierra Leone were the only countries that were shown to have increased U5MR in their second DHS survey compared to their first, specifically from 70 to 96 deaths per thousand live deaths for Benin and from 140 to 156 deaths for Sierra Leone. Contrarily, Mali experienced by far the largest decline in U5MR, i.e. from 191 deaths per thousand live births in 2006 to 95 deaths in 2012. The first in time DHS surveys were conducted between 2003 (Burkina Faso) and 2013 (Togo), while the surveys of the second set between 2010 (Burkina Faso) and 2017 (Benin and Togo).

Results from the bivariate geostatistical variable selection presented useful patterns of which variables are most important for subsequent model fit. Preventative health interventions belonging to either the category of vaccinations or reproductive health were most frequently associated with all-cause under-five mortality. Specifically, Bacillus Calmette–Guérin (BCG), DPT and measles immunization were found to be statistically important in all 21 countries. Polio and Tetanus vaccinations were shown to be somewhat less associated with under-five mortality, as we found 19 and 16 important effects on mortality respectively. Reproductive health, taking the form of either antenatal care (ANC) for pregnant mothers from a skilled provider or frequent, at least four, ANC visiting during pregnancy, had a significant impact on pre-school mortality in at least 18 countries. Additionally, vitamin A supplementation and

deworming medication provided to children over the past 6 months that the DHS surveys took place, were almost everywhere associated with mortality.

Our model-based estimates showed that approximately one out of three health interventions considered in this modelling study were associated with the changes in all-cause under-five mortality risk for at least one country. From all preventative or curative health interventions, the coverage of BCG vaccination was associated to mortality in 18 out of 21 countries, with the only exceptions of Ethiopia, Senegal and Togo. BCG immunization additionally had the strongest impact on the changes of mortality risk, in terms of posterior median estimates. Vitamin A supplementation followed BCG as it was found to be associated with changes in mortality in half of the countries. Deworming was statistically important in Benin, Burundi, Cameroon, Guinea, Malawi and Uganda. Sporadic associations between a small set of health interventions and pre-school mortality were found in a few countries, specifically for the following interventions: measles (Ghana, Mozambique and Zambia), immediate breastfeeding after birth for infants (Namibia and Nigeria), exclusive breastfeeding for children up to six months (Zambia), improved sanitation facilities in households (Togo) and use of ITNs from under-five children (Malawi).

5.4 Discussion

To our knowledge, this is the first study to assess the effect of preventive or curative health interventions on the changes in under-five mortality risk for 21 SSA countries, using routinely collected household survey data from the DHS program. We employed rigorous Bayesian variable selection procedures and fitted Bayesian geostatistical Weibull survival models to conduct our analysis. We included in the analysis countries with at least two DHS surveys that had available information on mortality, maternal, socio-economic and individual child characteristics.

Our results highlight that BCG immunization, vitamin A supplementation and deworming medication are the interventions mostly associated with the changes in all-cause under-five mortality risk in sub-Saharan Africa. To role of vaccinations in Africa, for which BCG

immunization constitutes an important pillar, has been well documented in the literature. Reduced under-five mortality risk in areas of complete vaccination coverage, compared to regions with no immunization, was reported from a modelling study utilizing DHS data from 149 countries (McGovern, 2015). The Global Alliance for Vaccines and Immunization (GAVI) reported that for the period 2000-2013, during which GAVI efforts helped the immunization of 440 million children in low-income countries, six million deaths were averted due to vaccinations (Bustreo, 2015). In terms of the role of Vitamin A, a pooled analysis (Stevens, 2015) of population-based surveys for 138 low-income and middle-income countries depicted that the African region is the most afflicted area globally in terms of vitamin A deficiencies, which in some countries like Mali could exceed 60%. Our analysis supports these results and demonstrates the crucial role that vitamin A supplementation can play in reducing under-five mortality risk.

We assumed that the under-five mortality for the first survey period could be solely explained by environmental and climatic factors. This assumption also affected the prediction of the mortality risk, in terms of scale parameter for the Weibull lifetime distribution, at locations of the second survey for the first survey year. This hypothesis was because we did not have available DHS data at locations of the second surveys for the first time point. Consecutively, while fitting using more confounders could have been possible, we would not have available data for prediction. Contrarily, climatic and environmental data are available across all regions for all years and thus extraction of such information at locations of the second survey for the first time point was possible. Additionally, environmental and climatic data can be used as proxies of infectious and other related diseases and in doing so they account for some of the leading causes of mortality in sub-Saharan Africa.

Our work uses statistical models to associate variables with mortality and does not imply causal relations. Instead, to establish such associations, a causal inference approach should have been rather used. The mortality data have been studied for a period of five years preceding the survey and our work did not take into account changes in the coverage of the health interventions over that period. This assumption might have resulted on

underestimation of the association between health interventions and changes in mortality risk.

Acknowledgments

The authors are grateful to the DHS for making the data available and all collaborating institutions whose contributions have enabled this study. The Swiss National Science Foundation (SNSF) Swiss Programme for Research on Global Issues for Development (R4D) project no. IZ01Z0-147286 funded this work.

Ethical approval and consent to participate

In this study we analysed secondary data made available by the Demographic Health Survey (DHS) MEASURE. According to survey protocols and related documents of the surveys, ethical approval was obtained from the Institutional Review Board of International Consulting Firm (ICF) of Calverton, Maryland, USA, and from the national ethical committees in the countries that the surveys were contacted. Details of ethical clearance are published in the DHS reports available at https://dhsprogram.com/publications/index.cfm.

Competing interests

We declare no competing interests.

Role of funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the paper. The authors had full access to all data in the study and had final responsibility for the decision to submit for publication.

Contributors

IP conceptualised the project, processed and analysed the data, interpreted the results, and wrote the manuscript. PV acquired the financial support of the project, conceptualised the project and assisted in statistical analysis. JU and PV revised the manuscript and provided important intellectual content. All authors approved the final version of the manuscript.

Availability of Materials and Data

The study data are available upon request from the Demographic and Health Surveys program (https://dhsprogram.com/data/Access-Instructions.cfm). The “Dataset Terms of Use” does not allow us to pass on downloaded data.

Table 5.1: Descriptive analysis of the data in 21 sub-Saharan African countries, collected from the DHS programme.

	1st survey				2nd survey			
	Survey year	Number of locations	Number of children	U5M	Survey year	Number of locations	Number of children	U5M
Benin	2011	746	13336	70	2017	540	13262	96
Burkina F.	2003	397	10575	184	2010	541	14159	129
Burundi	2010	376	7742	96	2016	552	13135	78
Cameroon	2004	461	8085	144	2011	577	11723	122
Ethiopia	2011	571	11253	88	2016	622	10221	67
Ghana	2008	401	2942	80	2014	423	5801	60
Guinea	2005	291	6245	163	2012	300	7039	123
Liberia	2007	291	5704	110	2013	322	7606	94
Malawi	2010	827	19490	112	2015	850	17286	64
Mali	2006	405	14194	191	2012	413	10326	95
Mozambique	2003	11	10326	152	2011	609	11089	97
Namibia	2006	486	5071	69	2013	537	5046	54
Nigeria	2008	886	28647	157	2013	889	31225	128
Rwanda	2010	492	9002	76	2014	492	7856	50
Senegal	2010	385	12103	72	2016	214	6725	51
Sierra Leone	2008	349	5590	140	2013	435	11938	156
Tanzania	2010	458	7750	81	2015	608	10233	67
Togo	2013	329	6979	88	2017	171	3415	73
Uganda	2011	400	7795	90	2016	685	15276	64
Zambia	2007	319	6401	119	2013	719	13412	75
Zimbabwe	2010	393	5372	84	2015	399	6135	69

Burkina F. stands for Burkina Faso.

Table 5.2: Description of health interventions.

Intervention	Description of Intervention
Malaria bednets	
%H_1ITN	Percentage of households with at least one Insecticide-Treated net (ITN)
%H_1ITN2	Percentage of households with at least one ITN for every two people
%P_ITNA	Percentage of population with access to an ITN within their household
%P_ITNS	Percentage of population in a household that slept under an ITN the previous night before the survey
%P_ITN5	Percentage of under-five children in a household who slept under an ITN the night before the survey
%P_ITNU	Percentage of existing ITNs used by the population in a household the previous night of the survey
WASH	
Water	Percentage of households with improved source of drinking water
Sanitation	Percentage of households with improved sanitation facilities
Reproductive health	
ANC prov.	Percentage of pregnant mothers receiving antenatal care (ANC) from a skilled provider
4+ ANC	Percentage of pregnant woman making at least four ANC visits during their pregnancy
Fansidar	Percentage of pregnant woman which received Fansidar during pregnancy
Post. Care	Percentage of new-borns receiving first postnatal check-up from a skilled provider within hours after delivery
Breastfeeding	
Immediate	Percentage of infants which breastfed within one hour after birth
Exclusive	Percentage of infants which exclusively breastfed during the first six months after delivery
Vaccinations	
Tetanus	Percentage of last-born children protected against neonatal Tetanus
BCG	Percentage of children vaccinated against Bacillus Calmette–Guérin (BCG)
DPT	Percentage of children with complete Diphtheria, Tetanus, and Pertussis (DPT) vaccination
Polio	Percentage of children with complete Polio vaccination
Measles	Percentage of children vaccinated against Measles
Micronutrients	
Vitamin A	Percentage of children which received vitamin A supplements in the past 6 months
Iron	Percentage of children which received iron supplements in the past 7 days
Iodized salt	Percentage of children which lived in a household with positive test for iodized salt
Treatments	
ORS	Percentage of children with diarrhoea given fluid from oral rehydration solution (ORS)
ACT	Percentage of children with fever during the two weeks prior to the survey which took Artemisinin-combination therapy (ACT)
Deworming	Percentage of children which took deworming medication in the past 6 months

Table 5.3: Effect of health interventions on changes in under-five mortality risk across 21 countries in sub-Saharan Africa.

	Benin	Burkina Faso	Burundi	Cameroon	Ethiopia	Ghana	Guinea
H_1ITN	X[1]	X	X	X	X	X	
H_1ITN2	X	X	X	X	X	X	X
P_ITNA	X	X	X	X	X	X	
P_ITNS	X	X	X	X	X	X	X
P_ITN5	X	X	X	X	X		X
P_ITNU	X	X	X	X	X	X	X
Water	X		X	X	X	X	X
Sanitation				X		X	
ANC prov.							
4+ ANC			X				
Fansidar		X	X	X	X		
Post. Care	X		X	X	X	X	X
Immediate	X	X	X			X	X
Exclusive	X	X	X	X	X	X	X
Tetanus			X				
BCG	-1.44 (-2.47,-0.41)	-3.75 (-4.80,-2.74)	-3.36 (-4.98,-1.76)	-2.95 (-3.84, -2.09)		-5.30 (-6.89,-3.75)	-2.15 (-3.15,-1.23)
DPT							
Polio							
Measles						-1.55 (-2.88,-0.20)	
Vitamin A	-0.55 (-0.98,-0.13)	-0.45 (-0.82,-0.06)	-0.73 (-1.36,-0.10)		-0.62 (-1.23,-0.01)		
Iron		X	X		X	X	X
Iodized Salt		X	X	X	X	X	X
ORS	X	X	X		X	X	
ACT	X	X	X	X	X		X
Deworming	-0.67 (-1.34,-0.03)		-1.02 (-1.72,-0.33)	-0.73 (-1.32,-0.15)			-1.27 (-2.21,-0.34)

1. Not selected for model fit, based on a bivariate geostatistical variable selection procedure.

	Liberia	Malawi	Mali	Mozambique	Namibia	Nigeria	Rwanda
H_1ITN	X	X	X	X	X	X	
H_1ITN2	X	X	X	X	X	X	X
P_ITNA	X	X	X	X	X	X	
P_ITNS	X			X	X	X	X
P_ITN5	X	-0.93 (-1.77,-0.10)		X	X	X	X
P_ITNU	X		X	X	X	X	X
Water	X	X	X	X	X		X
Sanitation	X	X	X	X	X		X
ANC prov.							
4+ ANC							
Fansidar	X		X		X		X
Post. Care	X	X	X	X	X		X
Immediate	X		X		-0.87 (-1.59,-0.15)	-0.35 (-0.64,-0.05)	
Exclusive	X	X	X	X	X	X	X
Tetanus					X		X
BCG	-2.50 (-3.58,-1.44)	-3.73 (-5.06,-2.42)	-1.45 (-2.37,-0.55)	-2.79 (-3.68,-1.95)	-5.54 (-7.03,-4.10)	-0.91 (-1.52,-0.31)	-8.65 (-11.29,-6.03)
DPT							
Polio							
Measles				-1.32 (-2.17,-0.45)			
Vitamin A		-0.89 (-1.54,-0.27)	-0.70 (-1.30,-0.09)	-0.71 (-1.40,-0.03)	-1.09 (-2.02,-0.14)	-0.56 (-0.99,-0.12)	
Iron		X			X	X	X
Iodized Salt	X		X	X	X	X	X
ORS	X	X	X		X	X	X
ACT	X	X	X	X	X	X	X
Deworming		-0.91 (-1.55,-0.25)					

	Senegal	Sierra Leone	Tanzania	Togo	Uganda	Zambia	Zimbabwe
H_1ITN	X	X	X	X		X	X
H_1ITN2	X	X	X	X		X	X
P_ITNA	X	X	X	X		X	X
P_ITNS	X	X	X	X			X
P_ITN5	X	X	X	X			X
P_ITNU		X	X	X	X	X	X
Water	X	X	X		X	X	X
Sanitation		X	X	-2.41 (-5.12,-0.02)	X	X	X
ANC prov.			X				
4+ ANC			X			X	
Fansidar		X	X				X
Post. Care		X	X	X	X	X	X
Immediate	X	X	X		X		
Exclusive	X	X	X	X	X	-2.90 (-4.31,-1.53)	
Tetanus			X			X	
BCG		-4.04 (-5.15,-2.90)	-2.71 (-4.19,-1.28)		-2.45 (-3.60,-1.36)	-4.80 (-6.01,-3.62)	-2.27 (-4.11,-0.43)
DPT							
Polio			X				X
Measles						-1.27 (-2.19,-0.34)	
Vitamin A		-0.90 (-1.58,-0.22)					
Iron	X		X	X	X	X	X
Iodized Salt	X		X	X		X	X
ORS	X	X		X	X	X	X
ACT	X	X	X	X	X	X	X
Deworming					-0.90 (-1.47,-0.33)		X

5.5 Appendix

Model specification

The bivariate geostatistical variable selection, the model fitting of the climatic data for the first survey year and the subsequent prediction at the locations of the second survey were all conducted in INLA, while the final model fit for the effect of health interventions on changes in under-five mortality risk was performed in STAN. The two platforms follow a different parametrization for the Weibull survival distribution that we assumed for our mortality data. Specifically, while both parametrizations consider a fixed shape parameter α for the Weibull distribution, INLA models a scale parameter lambda and STAN models a scale parameter sigma that are connected through the following equation:

$$\lambda = \left(\frac{1}{\sigma}\right)^{\alpha}.$$

That means that based on the choice of the scale parameter, i.e. lambda or sigma, the probability density function, the hazard function and all related equations of the Weibull survival distribution changed accordingly. Below, we provide details about the model formulation under both software and we provide further information on the 3-step process model fit to derive the changes in mortality specification.

The Weibull formulation under INLA assumes that the scale parameter λ is reparametrized in terms of model parameters. We assumed that T_{ji} represents the age of a child, with j indicating each child and i corresponding to one of the n surveyed locations for which we had available data, i.e. $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$. Tau followed a Weibull lifetime distribution with fixed shape parameter alpha and scale parameter lambda that was linked with the linear predictor as follows:

$$\lambda_{ji} = \exp(\eta_{ji}),$$

$$\eta_{ji} = \beta_0 + \sum_{\kappa=1}^{K} \beta_\kappa x_{ji\kappa} + \xi_i.$$

To this equation, eta represented the linear predictor and the set of betas $\beta = (\beta_0, \beta_1, \dots, \beta_\kappa)^T$ indicated the K regression coefficients assigned to each predictor $x_{ji\kappa}$, with κ representing the regression coefficient, j each child and i each location. The random vector $\xi(s) = (\xi_1, \xi_2, \dots, \xi_n)^\tau$ corresponded to the value of geostatistical intercept at each location (cluster) at which we had available data. $\xi(s)$ followed a zero-mean, multivariate Gaussian distribution with Matérn covariance matrix, i.e. $\xi(s) \sim N(0, \Sigma_1)$ and $\Sigma_1(s_1, s_2) = \frac{\sigma_1^2(\kappa_1\, d(s_1,s_2))^\nu K_\nu(\kappa_1\, d(s_1,s_2))}{\Gamma(\nu)\, 2^{\nu-1}}$, with s_1 and s_2 indicating two different locations, $d(s_1, s_2)$ being the distance between s_1 and s_2, κ_1 being the scaling parameter and σ_1^2 being the spatial process variance. Additionally, K_ν was the modified Bessel function of second kind and order ν. The specification of the covariance matrix was completed with the spatial range parameter r_1, which under the equation $r_1 = \frac{\sqrt{8}}{\kappa_1}$ particularly determines the distance at which the spatial correlation becomes negligible. In the case of bivariate variable selection, $K = 1$ and we fitted each time, separately, one of the 25 health interventions considered in this study. In the case of the 3-step model fit, K was equal to the number of selected health interventions plus all confounders and differences in climate.

We compiled two different DHS surveys for each country and wanted to model the difference in mortality risk between the two time points. As the two different datasets derived from the surveys followed their own Weibull survival distribution, it followed that:

$$\lambda_1 = \exp(\eta_1),$$

$$\lambda_2 = \exp(\eta_2),$$

with λ_1 and λ_2 representing the scale parameter of the first and second DHS survey respectively and eta being the corresponding linear predictor. By taking the logarithmic difference between the two equations, it followed that:

$$\log(\lambda_2) - \log(\lambda_1) = \eta_2 - \eta_1 \implies \log(\lambda_2) = \log(\lambda_1) + \eta_2 - \eta_1 \implies$$

$$\lambda_2 = \lambda_1 \exp(\eta_2 - \eta_1).$$

To move towards the STAN specification we replaced λ_2 with its equivalent in STAN parametrization, i.e. $\left(\frac{1}{\sigma_2}\right)^{\alpha_2}$ and solved towards σ_2 in order to be able to get the equation needed for the STAN implementation:

$$\lambda_2 = \lambda_1 \exp(\eta_2 - \eta_1) \Longrightarrow \left(\frac{1}{\sigma_2}\right)^{\alpha_2} = \lambda_1 \exp(\eta_2 - \eta_1) \Longrightarrow \alpha_2 \log\left(\frac{1}{\sigma_2}\right) = \log(\lambda_1) + \eta_2 - \eta_1 \Longrightarrow$$

$$-\alpha_2 \log(\sigma_2) = \log(\lambda_1) + \eta_2 - \eta_1 \Longrightarrow \log(\sigma_2) = -\frac{\log(\lambda_1)}{\alpha_2} + \frac{\eta_2 - \eta_1}{-\alpha_2} \Longrightarrow \sigma_2$$

$$= \exp\left(\log\left(\lambda_1^{-\frac{1}{\alpha_2}}\right)\right)\exp\left(-\frac{\eta_2 - \eta_1}{\alpha_2}\right) \Longrightarrow$$

$$\sigma_2 = \lambda_1^{-\frac{1}{\alpha_2}} \exp\left(-\frac{\eta_2 - \eta_1}{\alpha_2}\right).$$

The last formula is what we explicitly modeled in STAN and consisted the last step (the 3rd step) of what we call a 3-step model fit. The first two steps of the model fit involved the spatial alignment of locations, as for a single country, the survey locations of the two DHS surveys were generally different. Therefore, in order to fit the last formula provided above in STAN, we had to fit (step1) λ_1 at the locations of the first DHS survey and subsequently predict (step2) λ_1 at the locations of the second DHS survey. Assuming that $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$ are the locations of the first DHS survey and $s^* = \{s_1^*, s_2^*, \dots, s_n^*\},\ s_i^* \subset R^2$ are the locations of the second DHS survey that are generally different than s, the first step involved the model fitting for the first survey in INLA as described above, i.e.

$$\lambda_1(s_i) = \exp\big(\eta_1(s_i)\big).$$

The second step involved the alignment of λ_1 at the second survey locations s^* and therefore we derived the distribution $\lambda_1(s_i^*)$ conditionally on the spatial process and regression parameters as follows:

$$\eta_1(s_i^*)/\beta_1, \xi_1(s)\ \sim N(\beta_1 X_1(s^*) + \Sigma_{s^*s}\Sigma_{ss}^{-1}\xi_1(s)\ , \Sigma_{s^*s^*} - \Sigma_{s^*s}\Sigma_{ss}^{-1}\Sigma_{ss^*})$$

$$\lambda_1(s_i^*) = \exp\big(\eta_1(s_i^*)\big).$$

In the above equations, β_1 represents the estimated regression coefficients, $X_1(s^*)$ is the set of climatic factors utilized in our analysis that were extracted at the locations of the latest

survey, $\xi_1(s)$ is the spatial process and Σ represents the covariance matrix evaluated at the different set of locations s and s^*. After having aligned (predicted) λ_1 at the second survey locations, we fitted the 3rd step of our analysis

$$\sigma_2(s_i^*) = \lambda_1^{-\frac{1}{\alpha_2}}(s_i^*) \exp\left(-\frac{\eta_2(s_i^*) - \eta_1(s_i^*)}{\alpha_2}\right).$$

Environmental and climatic data

Table 5.4: Sources of environmental and climatic data.

Data	Source	Spatial resolution
Annual average Normalised Difference Vegetation Index (NDVI)	MODIS	1x1 km^2
Annual average Day and Night Land Surface Temperature (LST)	MODIS	1x1 km^2
Land Cover Type (LC)	MODIS	0·5x0·5 km^2
Distance from water bodies (DWATER)	MODIS	0·5x0·5 km^2
Annual average Rainfall	USGSS	8x8 km^2
Altitude (Digital Elevation model)	SRTM	0·5x0·5 km^2
Urban rural extent	GRUMP	1x1 km^2

MODIS: Moderate Resolution Imaging Spectroradiometer; USGSS: U.S. Geological Survey-Earth Resources Observation Systems; SRTM: Shuttle Radar Topographic Mission; GRUMP: Global Rural and Urban Mapping project

Additional Results

Table 5.5: Posterior estimates obtained from the bivariate geostatistical variable selection for the second DHS survey data of each country. We present posterior summaries (posterior median, 95% Bayesian Credible Interval) only for the statistically important associations between health interventions and under-five mortality.

	Benin	Burkina Faso	Burundi	Cameroon	Ethiopia	Ghana	Guinea
H_1ITN							-0.66 (-1.20,-0.13)
H_1ITN2							
P_ITNA							-0.80 (-1.58,-0.02)
P_ITNS							
P_ITN5						-0.73 (-1.42,-0.03)	
P_ITNU							
Water		-0.47 (-0.74,-0.20)					
Sanitation	-1.00 (-1.57,-0.42)	-0.74 (-1.19,-0.29)	-0.70 (-1.07,-0.34)		-1.11 (-1.84,-0.37)		-0.89 (-1.40,-0.39)
ANC prov.	-0.90 (-1.39,-0.42)	-0.68 (-1.01,-0.34)	-2.22 (-3.25,-1.19)	-1.12 (-1.63,-0.61)	-0.97 (-1.38,-0.56)	-2.42 (-3.57,-1.27)	-0.99 (-1.43,-0.54)
4+ ANC	-0.96 (-1.47,-0.46)	-0.53 (-1.05,-0.01)		-1.30 (-1.78,-0.82)	-1.29 (-1.79,-0.79)	-1.27 (-2.11,-0.42)	-1.08 (-1.61,-0.55)
Fansidar	-1.10 (-1.58,-0.62)					-1.87 (-2.81,-0.94)	-0.61 (-1.20,-0.03)
Post. Care		-0.65 (-1.21,-0.09)				-1.00 (-1.96,-0.04)	
Immediate				-0.38 (-0.77,0)	-0.58 (-1.03,-0.13)	-0.65 (-1.28,-0.03)	
Exclusive							
Tetanus	-1.13 (-1.62,-0.64)	-1.80 (-2.39,-1.21)		-1.30 (-1.86,-0.75)	-1.15 (-1.62,-0.68)	-1.72 (-2.74,-0.69)	-1.30 (-1.84,-0.76)
BCG	-1.94 (-2.51,-1.38)	-3.09 (-3.52,-2.66)	-3.91 (-4.77,-3.04)	-2.85 (-3.29,-2.42)	-1.49 (-2.02,-0.96)	-4.80 (-5.85,-3.76)	-1.75 (-2.19,-1.30)
DPT	-1.61 (-2.15,-1.08)	-2.02 (-2.40,-1.63)	-3.56 (-4.40,-2.72)	-1.64 (-2.00,-1.30)	-1.36 (-1.91,-0.82)	-2.43 (-3.29,-1.58)	-0.93 (-1.36,-0.49)
Polio	-1.43 (-2.00,-0.87)	-1.77 (-2.13,-1.41)	-2.47 (-3.20,-1.74)	-1.30 (-1.67,-0.93)	-2.13 (-2.93,-1.33)	-1.33 (-2.04,-0.64)	-1.13 (-1.61,-0.66)
Measles	-1.52 (-2.15,-0.89)	-2.39 (-2.81,-1.96)	-2.87 (-3.67,-2.08)	-2.19 (-2.61,-1.76)	-1.57 (-2.21,-0.92)	-3.61 (-4.62,-2.60)	-1.33 (-1.80,-0.87)
Vitamin A	-0.74 (-1.09,-0.39)	-1.00 (-1.31,-0.70)	-1.55 (-2.02,-1.08)	-0.88 (-1.32,-0.45)	-1.20 (-1.66,-0.75)	-1.13 (-1.83,-0.44)	-0.79 (-1.25,-0.32)
Iron	-0.48 (-0.93,-0.02)			-1.40 (-2.38,-0.42)			
Iodized Salt	-0.68 (-1.32,-0.05)						
ORS				-1.80 (-3.48,-0.12)			-1.78 (-3.42,-0.14)
ACT						-0.61 (-1.02,-0.19)	
Deworming	-0.83 (-1.27,-0.39)	-0.50 (-1.00,-0.01)	-1.59 (-2.11,-1.08)	-1.72 (-2.17,-1.28)	-1.77 (-2.62,-0.91)	-1.43 (-2.29,-0.59)	-1.65 (-2.34,-0.95)

	Liberia	Malawi	Mali	Mozambique	Namibia	Nigeria	Rwanda
H_1ITN							-0.77 (-1.53,-0.02)
H_1ITN2							
P_ITNA							-0.72 (-1.45,0)
P_ITNS		-0.59 (-1.06,-0.13)	-0.96 (-1.68,-0.25)				
P_ITN5		-0.72 (-1.12,-0.33)	-0.93 (-1.53,-0.33)				
P_ITNU		-0.56 (-1.01,-0.12)					
Water						-0.34 (-0.48,-0.20)	
Sanitation						-0.41 (-0.61,-0.22)	
ANC prov.	-1.26 (-2.07,-0.45)	-1.99 (-2.71,-1.27)	-1.23 (-1.72,-0.74)	-0.94 (-1.59,-0.28)	-1.45 (-2.52,-0.38)	-0.83 (-1.10,-0.56)	-1.99 (-3.36,-0.63)
4+ ANC	-0.78 (-1.38,-0.18)	-0.90 (-1.45,-0.35)	-1.51 (-2.13,-0.88)	-0.78 (-1.35,-0.21)	-1.01 (-1.85,-0.16)	-0.80 (-1.09,-0.50)	-1.50 (-2.38,-0.62)
Fansidar		-1.94 (-2.63,-1.25)		-0.98 (-1.54,-0.42)		-0.93 (-1.35,-0.51)	
Post. Care						-1.72 (-2.34,-1.09)	
Immediate		-0.75 (-1.23,-0.27)		-0.62 (-1.09,-0.14)	-0.90 (-1.66,-0.13)	-0.71 (-0.98,-0.43)	-1.45 (-2.33,-0.56)
Exclusive							
Tetanus	-1.15 (-1.82,-0.48)	-1.56 (-2.19,-0.93)	-1.08 (-1.64,-0.52)	-0.81 (-1.38,-0.24)		-0.88 (-1.16,-0.59)	
BCG	-1.94 (-2.46,-1.42)	-2.79 (-3.52,-2.06)	-1.40 (-1.82,-0.98)	-3.08 (-3.61,-2.55)	-5.89 (-7.01,-4.77)	-1.11 (-1.33,-0.89)	-8.26 (-9.56,-6.96)
DPT	-1.14 (-1.63,-0.65)	-2.18 (-2.89,-1.47)	-1.16 (-1.59,-0.73)	-1.83 (-2.28,-1.40)	-2.48 (-3.30,-1.66)	-1.30 (-1.56,-1.05)	-5.44 (-6.43,-4.44)
Polio	-0.79 (-1.32,-0.26)	-0.99 (-1.59,-0.39)	-0.96 (-1.52,-0.41)	-1.44 (-1.85,-1.02)	-1.59 (-2.36,-0.81)	-1.40 (-1.73,-1.07)	-2.81 (-3.53,-2.09)
Measles	-1.83 (-2.40,-1.26)	-1.30 (-2.02,-0.58)	-1.30 (-1.77,-0.83)	-2.59 (-3.09,-2.09)	-2.69 (-3.54,-1.83)	-1.25 (-1.51,-0.99)	-0.74 (-1.34,-0.14)
Vitamin A	-0.75 (-1.20,-0.30)	-1.53 (-2.00,-1.06)	-1.21 (-1.63,-0.79)	-2.00 (-2.48,-1.53)	-2.83 (-3.58,-2.08)	-1.11 (-1.38,-0.84)	-3.28 (-4.14,-2.42)
Iron	-0.70 (-1.39,-0.01)		-1.46 (-2.06,-0.86)	-0.72 (-1.27,-0.17)			
Iodized Salt		-0.40 (-0.71,-0.09)					
ORS				-1.44 (-2.72,-0.15)			
ACT							
Deworming	-0.78 (-1.33,-0.22)	-1.48 (-1.97,-1.00)	-0.91 (-1.43,-0.40)	-1.26 (-1.75,-0.78)	-1.07 (-1.77,-0.37)	-1.05 (-1.42,-0.68)	-3.20 (-4.07,-2.32)

	Senegal	Sierra Leone	Tanzania	Togo	Uganda	Zambia	Zimbabwe
H_1ITN					-0.95 (-1.45,-0.46)		
H_1ITN2					-0.70 (-1.15,-0.25)		
P_ITNA					-0.89 (-1.37,-0.42)		
P_ITNS					-0.69 (-1.15,-0.22)	-0.60 (-1.13,-0.06)	
P_ITN5					-0.59 (-1.01,-0.18)	-0.51 (-0.94,-0.08)	
P_ITNU	-0.69 (-1.28,-0.10)						
Water				-0.42 (-0.74,-0.09)			
Sanitation	-0.83 (-1.32,-0.34)			-1.66 (-2.69,-0.64)			
ANC prov.	-1.84 (-2.96,-0.72)	-1.02 (-1.91,-0.12)		-0.83 (-1.30,-0.37)	-1.40 (-2.19,-0.60)	-1.05 (-1.78,-0.33)	-2.38 (-3.10,-1.66)
4+ ANC	-2.35 (-3.26,-1.45)	-0.76 (-1.29,-0.22)		-1.09 (-1.76,-0.42)	-0.78 (-1.38,-0.18)		-1.81 (-2.60,-1.03)
Fansidar	-1.91 (-2.93,-0.89)			-1.60 (-2.22,-0.97)	-1.04 (-1.68,-0.40)	-0.92 (-1.57,-0.28)	
Post. Care	-1.91 (-2.65,-1.18)						
Immediate				-0.64 (-1.08,-0.20)		-0.48 (-0.91,-0.05)	-0.77 (-1.49,-0.06)
Exclusive						-1.76 (-3.03,-0.49)	-1.56 (-2.41,-0.70)
Tetanus	-1.69 (-2.76,-0.61)	-1.20 (-2.05,-0.35)		-1.47 (-2.23,-0.71)	-1.12 (-1.79,-0.46)		-2.23 (-3.02,-1.45)
BCG	-2.33 (-3.20,-1.46)	-3.46 (-4.05,-2.86)	-1.93 (-2.77,-1.08)	-3.80 (-4.54,-3.07)	-2.76 (-3.54,-1.98)	-4.58 (-5.25,-3.91)	-1.68 (-2.61,-0.75)
DPT	-2.13 (-2.86,-1.40)	-1.46 (-1.90,-1.03)	-1.13 (-1.88,-0.38)	-1.97 (-2.52,-1.43)	-2.16 (-2.82,-1.50)	-2.32 (-2.84,-1.80)	-1.06 (-2.01,-0.10)
Polio	-1.38 (-2.03,-0.73)	-1.41 (-1.84,-0.98)		-1.60 (-2.21,-0.99)	-1.81 (-2.46,-1.16)	-0.66 (-1.12,-0.20)	
Measles	-2.13 (-2.90,-1.37)	-2.22 (-2.79,-1.66)	-1.15 (-1.99,-0.31)	-2.18 (-2.85,-1.51)	-1.78 (-2.51,-1.06)	-3.42 (-4.02,-2.81)	-1.82 (-2.87,-0.78)
Vitamin A	-2.23 (-3.14,-1.33)	-1.77 (-2.25,-1.30)	-0.85 (-1.42,-0.29)	-2.63 (-3.22,-2.03)	-0.96 (-1.39,-0.52)	-2.05 (-2.55,-1.54)	-1.37 (-2.05,-0.69)
Iron		-0.90 (-1.39,-0.41)					
Iodized Salt		-0.40 (-0.72,-0.08)			-0.90 (-1.70,-0.11)		
ORS			-1.92 (-3.57,-0.28)				
ACT							
Deworming	-1.95 (-2.76,-1.15)	-0.94 (-1.45,-0.43)	-0.81 (-1.36,-0.27)	-0.68 (-1.19,-0.18)	-1.34 (-1.81,-0.87)	-1.66 (-2.16,-1.15)	

Chapter 6

Bayesian variable selection methods for spatially varying coefficient models

Isidoros Papaioannou[1,2], Jürg Utzinger[1,2], Penelope Vounatsou[1,2]

Affiliations:

[1]Swiss Tropical and Public Health Institute, Basel, Switzerland

[2]University of Basel, Basel, Switzerland

Abstract

In the African setting, control programmes and decision makers are interested in assessing the effects of health interventions on child mortality risk and identifying potential subnational differences of their effects. Geostatistical models provide a modelling framework for determining such associations and incorporating in the analysis distance-dependent spatial correlations between pairs of locations in which the observed data are measured. These models can be extended to include spatially varying coefficients, so that the dependency between the outcome and predictors is additionally studied over regions of interest. In the Bayesian approach, the conditional autoregressive (CAR) structure has been extensively used to model the covariance matrix of the spatially varying coefficients. As CAR-structured models are highly parameterized, variable selection procedures could be employed in order to reduce the parameter space, improve prediction accuracy and overall offer faster statistical analysis. Stochastic search variable selection (SSVS) augments the parameter space with latent inclusion-indicator variables and their posterior distribution is used for variable selection inference. For spatially varying coefficients, either global methods, suggesting the form of predictors across the whole spatial study area, or local methods exploring regional effects could be employed. We designed a simulation study and analysed data stemming from the Burundi 2016 Demographic and Health survey, in order to assess the performance of SSVS for geostatistical models and to compare three different selection methods for CAR-structured spatially varying coefficient models. We further studied the sensitivity of SSVS method when an additional selection procedure for varying coefficients is simultaneously employed. Our analysis showed that SSVS is able to accurately separate the statistically important predictors from the parameter space and is not influenced by the simultaneous implementation of another selection method. Results from the three spatially varying coefficient selection methods were relatively inconclusive**,** albeit one of them, the

Global selection method was able to correctly identify the true varying coefficients with 70 per cent success probability.

6.1 Introduction

In the era of Sustainable Development Goals (SDGs) under-five mortality still remains a highly relevant public health issue, with only 57 countries globally having met the Millennium Development Goals (MDGs). In Africa, despite large reductions in child mortality rates, it is estimated that multiple sites need an annual reduction of as much as 8.8%, so that the SDG 3.2 target for under-5 mortality is to be met (Golding, 2017). Relevantly, health interventions provide an important pillar against pre-school mortality (Shoo, 2007). In the absence of developed civil registration systems in most African countries, the Demographic and Health Surveys (DHS) have been developed and employed in order to measure and provide key information on health-related indicators, such as of child and maternal health, malaria, nutrition, household and other characteristics. Model-based studies are commonly utilizing the cross-sectional DHS datasets to associate the effects of health interventions and prevalence of diseases with child mortality and explore possible subnational variation of these effects (Papaioannou, 2019; Giardina, 2014; Millogo, 2019; Nambuusi, 2019). Given the wealth of the health indicators measured by data sources such as the DHS, variable selection techniques could be implemented to identify the most relevant indicators for subsequent model fitting.

Variable selection methods have been traditionally utilized to perform faster statistical analysis and improve prediction accuracy (Reunanen, 2003). This is achieved by selecting a subset of all predictors which are initially considered for model fit, based on some criteria that each variable selection method sets. In Bayesian analysis, most of the Bayesian variable selection (BVS) methods can be classified as either model choice-specific or variable-specific. The former refers to methods which use information criteria for model selection, while the latter refers to methods which support inclusion of variables based on

posterior inclusion probabilities (O'Hara, 2009). The variable-specific procedures can be implemented to both fixed and random components of a model (Frühwirth-Schnatter, 2010). Variable-specific selection methods can be classified into four categories (O'Hara, 2009): stochastic search variable selection (SSVS), indicator model selection, adaptive shrinkage and model space approach. SSVS (George, 1993) was the first BVS procedure to introduce the choice of variables for the fixed effects of a model. SSVS proposed to augment the parameter space of a model with latent variables which would act as indicators of whether a predictor should be included in a final model fit. Specifically, the latent variables were incorporated in the prior distribution of each regression coefficient and, based on their posterior marginal distribution, inference for selecting the corresponding predictor could be made. The term stochastic search refers to how the model space is explored (Peng, 2013). Extensions of SSVS were later proposed (Brown, 1998; George, 1997). Other notable stochastic search methods include: the Kuo and Mallick method (Kuo, 1998) (indicator model selection), the Gibbs variable selection (Dellaportas, 2002) (indicator model selection), Jeffreys' prior (Xu, 2003) (adaptive shrinkage) and the Reversible jump MCMC (Green, 1995) (Model space approach) method. A recent review[8] of numerous BVS procedures still argued for the usefulness of the SSVS method. Applications of stochastic search methods can be found, among others, in schistosomiasis (Chammartin, 2013) and malaria modelling (Giardina, 2014), Biology (García-Díaz, 2015), cancer research (Newcombe, 2014), genetics (Srivastana, 2009) and clinical research (Whitlock, 2014).
BVS procedures have also been developed for the random effects of a model (Kinney, 2007; Tüchler, 2008; Yang, 2011; Chen, 2003; Wagner, 2012). Spatially varying coefficient models constitute a subclass of all random effect models and their specification allows coefficients to vary in space, so that dependency between the outcome and predictors is better explained over a region of interest (Gelfand, 2003). One approach to BVS for this type of models, is the introduction of indicators which would specify one of the following forms for each predictor: exclusion of predictor, inclusion as fixed effect or inclusion as a spatially varying coefficient (Reich, 2010). The previous method can be characterized as "global" selection method,

since the proposed form of predictors is not determined at subregional (local) level but rather across the whole spatial study area. Contrarily, an approach for local BVS is to assume independent inclusion indicators for each coefficient and subregion, while imposing a spatial structure to the coefficient via a Gaussian copula (Boehm Vock, 2015). Another example of local BVS is to directly modify the covariance matrix of each regression coefficient, by left and right multiplication with a diagonal matrix, which holds an indicator vector in its main diagonal (Zhang, 2016). Recent efforts have introduced the spatial dependency directly to the selection indicators (Choi, 2016), for instance by imposing a spatial autologistic Besag model to each indicator of a coefficient and region.

Geostatistical models are a class of mixed models which assume spatial dependency in the response across a set of locations. Typically, a zero-mean Gaussian distributed random intercept is added to linear predictor, which assumes a spatially structured covariance matrix (Brown, 2015). A distance-dependent covariance matrix is usually considered. Additionally, in the case of geostatistical modelling with spatially varying coefficients, the vector of varying effects for a single coefficient usually follows a multivariate Normal distribution and is fully characterized by its covariance matrix (Finley, 2011). Among other specifications, conditional autoregression (CAR) modelling of the spatially varying components has been extensively used for spatial data analysis (Gelfand, 2003b). Despite the relatively recent proposals of new variable selection methods for spatially varying coefficient models, studies assessing the sensitivity of the methods under different assumptions are rather few.

The aim of this article is threefold: to assess the performance of SSVS for geostatistical models, to compare three different BVS methods for CAR-structured spatially varying coefficient models and to assess the sensitivity of SSVS when is co-implemented with a spatially varying coefficient selection method. In section 2 we provide an overview of the data, methodology and models utilized in this study. In section 3 we present details from the simulation study which we designed and conducted for all the proposed methods, while in section 4 we specify all relevant information from the application of the methods to the

Burundi DHS dataset. In section 5, we provide a detailed discussion of the simulation and application results.

6.2 Methods

Data sources

We extracted data stemming from the Burundi DHS which took place between October 2016 and March 2017. The rational of the survey was to provide a platform for measuring and updating key indicators on basic health and demographics. Data collection was based on a stratified, two-stage analysis and four different forms of questionnaires, i.e. household, individual woman, male and biomarker were utilized. We obtained data on individual child, maternal and household characteristics, malaria interventions, vaccinations, breastfeeding, macronutrients and reproductive health. The data were geolocated, i.e. latitude and longitude information was available for each observation. A detail summary of the survey methodology and data outcomes can be found on the Burundi DHS final report. Environmental and climatic data were downloaded from open access and remote sensing sources. We obtained from Moderate Resolution Imaging Spectroradiometer (MODIS) information on the land surface temperature (LST), the normalized difference vegetation index (NDVI) and the type of land cover (LC). We classified all observations to belong to a forest, grassland or cropland area based on the corresponding coordinates and LC type of a particular observation. The U.S. Geological Survey-Earth Resources Observation Systems (USGSS) Data Portal and the Shuttle Radar Topography Mission (SRTM) provided the portals through which rainfall and altitude data were correspondingly extracted.

Statistical models

Let $s = \{s_1, s_2, \dots, s_n\},\ s_i \subset R^2$ be a set of locations at which the outcome and predictors were observed. The response variable was defined as T_i and represented time to death for the i-th under-five child observed at the s_i location. Under the Weibull survival regression, the outcome T_i followed a Weibull distribution for which the shape parameter α was considered fixed and the scale λ was reparameterised in terms of predictors and regression coefficients.

By defining as $i = 1, \dots, N$ the number of observations, $p = 1, \dots, P$ the number of fixed effect parameters of the model and η_i the linear predictor corresponding to the i-th observation, the Weibull survival geostatistical model was formulated as:

$$T_i \sim Weibull(\alpha, \lambda_i),$$

$$\lambda_i = \exp(\eta_i),$$

$$\eta_i = \beta_0 + \sum_{p=1}^{P} \beta_p x_{ip} + w_{i0},$$

where β_0 is the constant term, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ is the vector of fixed effects, $\boldsymbol{X}_{NxP}$ the design matrix and $\boldsymbol{w}_0$ the geostatistical intercept. We used a N-dimensional, zero-mean Gaussian distribution $\boldsymbol{w}_0 \,/\, \boldsymbol{\Sigma}_0 \sim N_N(\mathbf{0}, \boldsymbol{\Sigma}_0)$ to model the location effects. The geostatistical intercept can be fully characterized by its covariance matrix with exponential correlation function $\boldsymbol{\Sigma}_0(s_i, s_j) = \sigma_0^2 exp(-\varphi \lambda_{ij})$, where σ_0^2 is the variance of the Gaussian process, φ the spatial decay parameter and $\boldsymbol{\Lambda}$ a distance matrix storing the kilometre distance λ_{ij} between s_i and s_j locations.

The above geostatistical model was further extended to include spatially varying effects. By assuming $k = 1, \dots, K$ varying coefficients at $m = 1, \dots, M$ locations, the linear predictor of the spatially varying coefficient model was defined as:

$$\eta_i = \beta_0 + \sum_{p=1}^{P} \beta_p x_{ip} + w_{i0} + \sum_{\substack{k=1 \\ i \in m}}^{K} D_{ik} w_{mk},$$

where $\boldsymbol{D}_{NxK}$ is the design matrix for the varying components of the model and w_{mk} represents the k-th varying coefficient at province m. The term $i \in m$ describes at which province the i-th observation belongs to and ensures that the $D_{ik} w_{mk}$ multiplication is compatible, as only one out of all M provinces corresponds to each observation. The varying effect $\boldsymbol{w}_k$, $k = 1, \dots, K$ is M-dimensional and represents the province-specific effect for covariate k. We could have obtained $K = P$ if all fixed effects P had a spatially varying component. The local varying effects were realizations of a Gaussian process, i.e. $w_k / \beta_k, \Sigma_\kappa \sim N_M(\boldsymbol{\beta}_\kappa, \Sigma_\kappa)$ with $\boldsymbol{\beta}_\kappa$ being a M-dimensional vector of the fixed effect β_k and Σ_κ the

corresponding covariance matrix. For the covariance matrix Σ_k we imposed a CAR structure, i.e. $\Sigma_k^{-1} = \sigma_k^{-2}(R - \rho_k\Omega)$, with R being a diagonal matrix with entries the sum of neighbours for each province and Ω a proximity matrix which reflected the neighbouring structure between the provinces.

We completed the Bayesian formulation with the assignment of suitable priors for each model parameter. We selected an inverse gamma $IG(2,1)$ distribution for all variance parameters and a gamma $G(1,1)$ for the shape of Weibull distribution. The spatial decay parameter phi followed a fragmented gamma distribution with shape and scale parameters equal to 1 and 0.5 respectively. Phi was restricted based on the rule that $\frac{-\log(0.05)}{\varphi}$ determines the distance at which correlation drops to 0.05 and it is considered as negligible. The spatial association measure of CAR model ρ_k ensures that $R - \rho_k\Omega$ is a non-singular matrix (Banerjee, 2014) and thus followed a uniform distribution, bounded by the inverse of the minimum and maximum of the ordered eigenvalues of the matrix $R^{-1/2}\Omega R^{-1/2}$.

Bayesian Variable Selection

Stochastic search variable selection was implemented by the introduction of a spike and slab prior for each regression coefficient β_p, $p = 1, \dots, P$, which means, in precision terms for the Normal distribution, that $\beta_p/\delta_p, \tau^2 \sim N(0, \delta_p\tau^2 + (1-\delta_p)u_0\tau^2)$. The indicator δ_p had a Bernoulli prior with 0.5 successes probability and dictated whether the corresponding coefficient β_p would be included in the model. When δ_p implied inclusion the prior distribution of β_p was transformed into $N(0, \tau^2)$ with an inverse gamma distribution imposed on the variance, while otherwise the shrinkage factor u_0 would concentrate the β_p prior around 0. We fixed u_0 equal to 0.001.

For the spatially varying coefficient models we considered a global BVS method, which determined whether a predictor should be included in a model and if so to either have a fixed or random slope. The global selection method introduced the vector indicator $\gamma_k = \binom{\gamma_{1k}}{\gamma_{2k}}$, which was sampled from a categorical distribution with three equally probable cases. These

three cases corresponded to exclusion of predictor from the model, inclusion as fixed effect and inclusion as a spatially varying coefficient, which translated into γ_k taking the form $\binom{0}{0}$, $\binom{1}{0}$ and $\binom{1}{1}$ respectively. The indicators γ_{1k} and γ_{2k} were incorporated as spike and slab priors in the regression coefficients β_k and the variance σ_k^2 of the CAR model, i.e. $\beta_k/\gamma_{1k}, \tau^2 \sim N(0, \gamma_{1k}\tau^2 + (1-\gamma_{1k})u_0\tau^2)$ and $\sigma_k^2/\gamma_{2k}, \widetilde{\sigma_k}^2 = \gamma_{2k}\widetilde{\sigma_k}^2 + (1-\gamma_{2k})u_0\widetilde{\sigma_k}^2$. We assigned 0 prior probability to the outcome $\binom{0}{1}$, as a predictor cannot have a spatially varying component without being included in the model at least as a fixed effect.

We also considered two alternatives to the global BVS method in order to allow variable selection indicators to vary locally in space. The main difference between global and local procedures was that local methods did not jointly define the spatially varying coefficient w_k with the aid of a single indicator γ_{2k}, but instead each of the elements of a varying coefficient (w_{mk}) was conditionally defined based on a province-specific indicator $\gamma_{mk}, m = 1, \dots, M$. Hence, under local selection strategies, the prior of the varying coefficient on the m-th province was $w_{mk}/\gamma_{mk}, \beta_k, \sigma_k^2 \sim N(\gamma_{mk}\beta_k, \gamma_{mk}\sigma_k^2 + (1-\gamma_{mk})u_0\sigma_k^2)$. Additionally, a Bernoulli distribution was assigned as prior on the indicators γ_{mk}. The two proposed local procedures differed on how the CAR model was imposed on the priors. The first strategy used a local success probability for the Bernoulli prior, that is $\gamma_{mk} \sim B(p_{mk})$, and adopted the probit link to impose a CAR structure on p_{mk}, i.e. $\Phi^{-1}(p_{mk}) = \varepsilon_{mk}$ and $\boldsymbol{\varepsilon}_k \sim N_M(\mathbf{0}, \boldsymbol{\Sigma}_k)$. The second strategy assigned a global probability of success for the Bernoulli prior, $p_k \sim U(0,1)$ and imposed a CAR model through a Gaussian copula on each w_{mk}. The copula specification was achieved by augmenting the parameter space with latent variables θ_{mk} and by utilizing the cumulative distribution function of the standard normal distribution, so that $\theta_k \,/\, \rho_k \sim N_M(\mathbf{0}, \boldsymbol{\Sigma}_k)$, with $\Sigma_k^{-1} = \sigma_k^{-2}(R - \rho_k\Omega)$ and $\theta_{mk} = \Phi^{-1}(F(w_{mk}))$. Since the copula model hypothesized on sampling from a great number of variables (133 in total) and in order to avoid overparametrization of the model, we relaxed the CAR specification by assuming $\sigma_k^2 = 1$.

A summary of all model formulations is provided in Table 6.1. The analysis was conducted in Just Another Gibbs Sampler (Plummer, 2003) (JAGS). Throughout the analysis we used 2 chains, which had thinning equal to 10, and we discarded initial iterations as burn-in which were equivalent to 5 per cent of the iterations we utilized in each chain. We selected the following number of updates for each case: 1 million for the simulation, 300 thousand for the application of SSVS and BVS on Burundi DHS and finally 100 thousand in order to get posterior summaries and compare between models for the selected covariates of our application. We assumed that mean posterior probability greater than 50 per cent favours inclusion of a predictor in a model or, when applicable, highlights an important varying effect for the specific province. Model comparisons were made based on the Watanabe-Akaike information criterion (Watanabe, 2010) (WAIC).

6.3 Simulation study

Overview

Our simulation study focused on the performance of: SSVS for geostatistical models, global and local BVS methods for spatially varying coefficient models and models combining the SSVS method with one of the BVS methods. To achieve that, we simulated 20 datasets based on a Weibull survival geostatistical model which was specified having $P = 10$ regressors, the first three of which were selected to be CAR-structured spatially varying. We hypothesized that our coefficients vary in an area of $M = 5$ polygons, e.g. provinces, and selected a neighbouring structure based on the following diagonal (R) and proximity (Ω) matrices:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{ and } = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

We assumed 500 observations of under-five children as outcome and georeferenced our data by assuming a sample size of ten at each location. The 50 unique locations used were

extracted from the Cote d'ivoire DHS 2011. Across all 20 simulations the values of the parameter space were held fixed and variability arose for each simulation from sampling for the following components: the $\boldsymbol{X}_{NxP}$ design matrix, the geostatistical term $\boldsymbol{w}_0$, the spatially varying parts w_{mk} and finally the generation of Weibull outcomes based on the linear predictor that we in each case had.

Our simulation study was designed so that only the first six regressors are statistically important, from which the first three vary in space. Yet, we tested the performance of the BVS methods assuming that the second, third and fourth coefficients were spatially varying. This specification is the common rule in practice as researchers do not have any prior knowledge about which regressors may vary in space and thus any reliable variable selection method should be able to separate true varying effects, which in our case translates into methods supporting inclusion as varying effect (global BVS) or associating important province effects (local BVS) only to the second and third regressors. Further details about the parameter values used to generate our simulations, the sampling distributions and the initial values we provided for sampling in JAGS are provided in the Appendix.

Results

Our simulation study indicated that the SSVS method was always able to identify correctly the statistically important coefficients and to support their inclusion in the model, with posterior mean probabilities exceeding 90 per cent. In geostatistical settings, only in one simulation study the SSVS wrongly proposed the inclusion of the last regression coefficient (β_{10}). When we added 3 spatially varying components to the geostatistical model (models 3 and 4) SSVS wrongly supported twice the inclusion of the last predictor and also in one occasion the inclusion of β_8. We obtained the same results when SSVS was coupled with the global BVS method. The addition of local BVS procedures to the stochastic search doubled the mistaken inclusion of β_8. A summary of results is presented in Table 6.2.

Global BVS method was able to identify correctly that all three regression coefficients β_2, β_3 and β_4 should be, at least as fixed effects, included in the model. It also hinted in seven out of ten times that β_2 is a coefficient that varies in space, while the results regarding the form of inclusion for β_3 and β_4 were inconclusive. Yet, by comparing the posterior summaries of the spatially varying effects for all 20 simulations for model 5, we were able to identify that in 14 out of 20 cases we found statistically important province effects for the true varying coefficients, i.e. β_2 and β_3, while not for β_4. When the global BVS was combined with a SSVS (model 6), the prior observation number was reduced to 12 cases.

Both local selection methods, which utilized the probit link and a Gaussian copula to impose a CAR spatial structure on the varying effects, failed to separate the true local varying coefficients and to correctly identify statistical important effects only on these regressors. By observing separately the outcomes of each simulation, in most occasions the local methods wrongly identify important local effects for β_4 which was designed to have only a fixed component. A summary of posterior inclusions for the spatially varying coefficients β_2, β_3 and β_4 and their local effects, stratified by variable selection, is presented in Table 6.3. Tables in Appendix summarize the results from the local BVS methods and specifically provide, for each simulation, all the indicators with posterior mean inclusion probability greater than 50 per cent.

6.4 Application

Overview

We applied SSVS and BVS procedures to select and analyse the effects of a set of child, maternal and household interventions on all-cause under-five mortality in Burundi. A description of the interventions coverage indicators can be found in Table 6.4. All models were adjusted for the cluster-level climatic conditions, as well as the rural-urban classification, at the geolocations of each observation, a socio-economic proxy based on the asset index of the family (classified as poorest or poorer), the sex, place of delivery and birth order for a child, the age at birth and live births of the mother as well as her educational

level. All intervention and adjustments were considered as fixed effects and thus the stochastic search analysed all of them. Additionally, in models 3 to 10 we imposed the CAR-structured spatially varying component to the following covariates: use of insecticide-treated nets (ITNs) by under-five children, antimalarial medication for children with fever and breastfeeding within 24 hours after birth.

Our application included 12’853 observations of under-five children across 540 unique locations, inside the 17 provinces of Burundi. An initial fit of all 10 models was performed in order to obtain the outcomes from the selection methods. We re-fitted all models based on the selected set of fixed effects and the type of inclusion that the BVS methods proposed for the varying variables. The latter meant for the local selection methods that in variables we found at least one local effect we re-fitted the model with a spatially varying component for the corresponding variable, otherwise it was considered as a fixed effect predictor. We estimated, for all the re-fitted models, the pointwise out-of-sample prediction accuracy using the WAIC. The goal of our application was not to analyse the effects of all health interventions on all-cause under-five mortality but rather utilize this publicly available dataset to encapsulate the advantages and drawbacks of the methods discuss in this article.

Results

Across all models imposing a SSVS method we found, similarly to our simulation, consistent results about which variables should be considered for the final model. Specifically, the use of ITNs from under-five children, antimalarial medication, breastfeeding of new-borns within 24 hours after birth, DPT vaccination, vitamin A supplementation, improved sanitation facilities, the birth order of a child, the number of live births from a mother and the asset index of a family were all highlighted for model inclusion with high posterior probabilities. Measles vaccination was the only predictor with mean values for the inclusion indicator just above 50 per cent and in one case (model 8) failed to surpass that threshold. A summary of results is presented in Table 6.5.

Greater variability of results was presented from the initial fit of the BVS methods. The use of ITNs from under-five children (β_2) was mostly proposed as a fixed effect, albeit the probit local selection method combined with SSVS (model 8) identified multiple important regional effects and thus implied inclusion of the predictor as a spatially varying coefficient. For antimalarial medication (β_3), all the local methods identified an important subregional effect in the Gitega province of Burundi. Contrarily, the Global BVS suggested that β_3 should be incorporated as either a fixed effect (model 6) or being excluded from the final model (model 5). For breastfeeding of new-borns, all models hinted inclusion as a fixed effect, except from model 6 which proposed the addition of a spatially varying component. A summary of results is presented in Table 6.6.

The initial analysis resulted on final models which were fitted to assess the effect of selected predictors on under-five mortality. Models A and B corresponded to the final form of the geostatistical and three-term (β_2, β_3 and β_4) spatially varying coefficient models. Model C derived from model 5 and included only fixed effects, with antimalarial medication not taken into account. Model D placed a varying coefficient on β_4 while model F on both β_2 and β_3. Models E and G had models 7 and 9–10 as their parent models and while both imposed a varying part on β_3, model E had one less fixed effect component (measles vaccination). Results across all final models showed that DPT3 vaccination, vitamin A supplementation, improved sanitation facilities and higher birth order bring about reduced mortality risk. Differently, poor asset index status for a family and many live births for a mother increase the mortality risk. The use of ITNs by under-five children showed relatively high negative mean posterior effects, but was identified as statistically important only in models C and F. Breastfeeding and measles vaccination were the only two predictors for which none of the models presented a statistically important association with the outcome and, from the set of variables included in the model, they were the two with the lowest inclusion posterior probabilities. Model D had the lowest WAIC and thus is seen as the best model fit, albeit

differences of WAIC between the models were relatively small and thus could be considered as negligible.

6.5 Discussion

To our knowledge, this is the first study to assess the performance of Bayesian variable selection methods for CAR-structured spatially varying coefficients, coupled with the stochastic search variable selection method. We also assessed the effects of specific child, maternal and household interventions on all-cause under-five mortality. Our simulation study revealed that SSVS is reliable for identifying suitable predictors from a candidate set, irrespective of the presence of other BVS methods. In terms of varying coefficients, the results were inconclusive about the usefulness of the methods, albeit we observed that the Global BVS method hints true varying coefficients with 70 per cent probability of success. DPT3 vaccination, vitamin A supplementation, improved sanitation facilities and higher birth order were shown to be associated with reduced mortality risk.

The utilization of a CAR parameterization with the additional parameter ρ_κ revealed that a lighter framework such as an exchangeable structure might be more appropriate, when the discussed methods are implemented. A modelling study (Boehm Vock, 2015) proposing the copula framework argued for the usefulness of the method, yet it proposed an exponential spatial correlation compared to the CAR structure. Similarly, a Bayesian probit model with regional selection (Zhang, 2016) tested its proposed method based on simulation studies which generating values, for the simulated regression coefficients, assuming either independence or an exponential correlation function.

Different assumptions about the choice of prior distributions and the corresponding values of the hyperparameters may have yielded dissimilar results. For the global selection method, we assumed a 3-dimensional categorical distribution for the inclusion indicator gamma, assigning a 34 per cent prior probability for the case of exclusion of the corresponding variable and 33 per cent probabilities for inclusion as fixed component or as varying coefficient. The aforementioned values hypothesized that the two forms of inclusion were

separate and equally probable, albeit together summed up to 66 per cent and thus left just 1 out of 3 chances for the exclusion case. In that sence, our choice of hyperparameters encouraged inclusion of variables. Another approach could have assumed 50 per cent prior probability for exclusion and 25 per cent for each of the two inclusion forms. Additionally, the categorical distribution could have been replaced by defining a joint prior for γ_{1k} and γ_{2k} which would rely on Beta distributed hyperprior probabilities (Reich, 2010). For the local selection method using the probit link, rather the logit link could have been employed. The use of inverse Gamma distributed priors for all variance terms could have been replaced by half-Cauchy, half-Normal or Uniform distributions, as supported by the corresponding literature (Gelman, 2006; Polson, 2012).

The selection of 50 per cent as an absolute cut-off of inclusion or exclusion of predictors from the linear part of the model, or as a measure which reveals which subregional effects were considered as statistically important, influenced our inference. Whereas very high or low mean posterior probabilities for the inclusion indicators lead to stronger conclusions, it is relatively unclear if variables with mean posterior indicators of 49 or 51 per cent should be explicitly removed or incorporated in a future model. For instance, SSVS showed that measles vaccination exceeded fractionally the cut-off, yet posterior summaries from the final models presented no association with the under-five mortality. Hence, a slightly harder inclusion cut-off of 55 per cent could have correctly prevented the use of measles vaccination from a predictor in the final models, assuming that the goal of inference from such methods is to result on a final model which incorporates only statistically important coefficients and facilitates a parameter space reduced to minimum. Contrarily, the delivery place of children achieved mean posterior values greater than 40 per cent, reaching in one occasion a value equal to 49, albeit the predictor was ultimately excluded from all final models. Analogously, the probit local selection method estimated a 49 per cent mean posterior value for the effect of breastfeeding in the Bururi province, but the effect was considered as not important.

The selection methods discussed in this study have also been extended to include temporal terms. The global selection method facilitated additional random temporal effects with the augmentation of the analogous set of inclusion indicators and forms through which a predictor was selected to be incorporated in the model (Cai, 2013). A modelling study associating the sea surface temperature, the latent heat flux and the occurrences of tropical cyclones (Cordoba, 2018) extended the Gaussian copula method to a spatio-temporal framework. Similarly, a spatially varying auto-regressive model was developed (Teng, 2018), to extent the idea of modelling brain voxels with homogeneous auto-regressive (AR) orders, with spatially varying AR orders which are additionally selected using a spike-and-slab prior at each voxel. The local BVS method assigning Bernoulli-prior indicators to local effects was recently enriched with Bernoulli distributed priors that are multiplied with temporally varying parameters of AR(1) structure (Choi, 2018).

The identification of health interventions and related risk factors which are associated with a decreased under-five mortality risk is crucial for decision makers. Additionally, the exploration of the geographical variation of such associations could assist in resource allocation, planning and implementation of health interventions. In our application we analysed a subset of all health interventions and verified their protective attribute across the whole spatial study area of Burundi. Routine intervention administration could benefit from variable selection schemes, such as those presented in our study, in an effort to identify areas in which successful implementation of health interventions has been achieved.

Ethical approval and consent to participate

In this study we analysed secondary data made available by the Demographic Health Survey (DHS) MEASURE. According to survey protocol and related documents of the survey, ethical approval was obtained from the Institutional Review Board of International Consulting Firm (ICF) of Calverton, Maryland, USA, and also from the national ethical committees in the country that the survey was conducted. Details of ethical clearance are published in the DHS reports available at https://dhsprogram.com/publications/index.cfm.

Acknowledgments

The authors are grateful to the DHS for making the data available and all collaborating institutions whose contributions have enabled this study. This work was funded by the European Research Council (ERC) Advanced Grant project no. 323180.

Competing interests

We declare no competing interests.

Role of funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the paper. The authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Contributors

IP processed and analysed the data, interpreted the results, and wrote the manuscript. PV conceptualised the project, assisted in statistical analysis, revised the manuscript and provided important intellectual content. All authors approved the final version of the manuscript.

Table 6.1: Summary of model formulations. Normal distribution is presented in precision terms.

Model	Fixed effects	Random effects	Inclusion indicators	Interpretation
1	$\beta_p \sim \mathcal{N}(0, 0.1),\ p = 1, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$	-	Geostatistical model
2	$\beta_p/\delta_p, \tau^2 \sim N(0, \delta_p \tau^2 + (1-\delta_p) u_0 \tau^2),$ $p = 1, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$	$\delta_p \sim \mathcal{B}(0.5),\ p = 1, \dots, 10$	Model 1 with SSVS imposed on fixed effects
3	$\beta_p \sim \mathcal{N}(0, 0.1),\ p = 1, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $\boldsymbol{w_k} / \beta_k, \sigma_k^2, \rho_k \sim \mathcal{N}_M(\boldsymbol{\beta_k}, \sigma_k^{-2}(\boldsymbol{R} - \rho_k \boldsymbol{\Omega})),\ k = 2,3,4$	-	Spatially varying coefficient model
4	$\beta_k \sim \mathcal{N}(0, 0.1),\ k = 2,3,4$ $\beta_p/\delta_p, \tau^2 \sim N(0, \delta_p \tau^2 + (1-\delta_p) u_0 \tau^2),$ $p = 1,5, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $\boldsymbol{w_k} / \beta_k, \sigma_k^2, \rho_k \sim \mathcal{N}_M(\boldsymbol{\beta_k}, \sigma_k^{-2}(\boldsymbol{R} - \rho_k \boldsymbol{\Omega})),\ k = 2,3,4$	$\delta_p \sim \mathcal{B}(0.5),\ p = 1,5, \dots, 10$	Model 3 with SSVS imposed on fixed effects
5	$\beta_p \sim \mathcal{N}(0, 0.1),\ p = 1,5, \dots, 10$ $\beta_k/\gamma_{1k}, \tau^2 \sim \mathcal{N}(0, \gamma_{1k} \tau^2 + (1-\gamma_{1k}) u_0 \tau^2),$ $k = 2,3,4$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $\boldsymbol{w_k} / \gamma_{2k}, \beta_k, \sigma_k^2, \rho_k \sim \mathcal{N}_M(\boldsymbol{\beta_k}, \sigma_k^{-2}(\boldsymbol{R} - \rho_k \boldsymbol{\Omega})),$ $\sigma_k^2/\gamma_{2k}, \widetilde{\sigma_k}^2 = \gamma_{2k} \widetilde{\sigma_k}^2 + (1-\gamma_{2k}) u_0 \widetilde{\sigma_k}^2,\ k = 2,3,4$	$\gamma = \binom{\gamma_{1k}}{\gamma_{2k}} \sim \mathcal{C}(0.34, 0.33, 0.33),\ k = 2,3,4$	Global BVS method for spatially varying coefficients
6	$\beta_p/\delta_p, \tau^2 \sim N(0, \delta_p \tau^2 + (1-\delta_p) u_0 \tau^2),$ $p = 1,5, \dots, 10$ $\beta_k/\gamma_{1k}, \tau^2 \sim \mathcal{N}(0, \gamma_{1k} \tau^2 + (1-\gamma_{1k}) u_0 \tau^2),$ $k = 2,3,4$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $\boldsymbol{w_k} / \gamma_{2k}, \beta_k, \sigma_k^2, \rho_k \sim \mathcal{N}_M(\boldsymbol{\beta_k}, \sigma_k^{-2}(\boldsymbol{R} - \rho_k \boldsymbol{\Omega})),$ $\sigma_k^2/\gamma_{2k}, \widetilde{\sigma_k}^2 = \gamma_{2k} \widetilde{\sigma_k}^2 + (1-\gamma_{2k}) u_0 \widetilde{\sigma_k}^2,\ k = 2,3,4$	$\gamma = \binom{\gamma_{1k}}{\gamma_{2k}} \sim \mathcal{C}(0.34, 0.33, 0.33),\ k = 2,3,4$ $\delta_p \sim \mathcal{B}(0.5),\ p = 1,5, \dots, 10$	Model 5 with SSVS imposed on fixed effects
7	$\beta_p \sim \mathcal{N}(0, 0.1),\ p = 1, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $w_{mk}/\gamma_{mk}, \beta_k, \sigma_k^2 \sim N(\gamma_{mk} \beta_k, \gamma_{mk} \sigma_k^2 + (1-\gamma_{mk}) u_0 \sigma_k^2),$ $k = 2,3,4,\ m = 1, \dots, 5$	$\gamma_{mk} / p_{mk} \sim \mathcal{B}(p_{mk}),$ $\Phi^{-1}(p_{mk}) = \varepsilon_{mk},$ $\boldsymbol{\varepsilon_k} / \sigma_k^2, \rho_k \sim \mathcal{N}_M(\mathbf{0}, \sigma_k^{-2}(\boldsymbol{R} - \rho_k \boldsymbol{\Omega})),$ $k = 2,3,4,\ m = 1, \dots, 5$	Local BVS method utilizing the probit link
8	$\beta_k \sim \mathcal{N}(0, 0.1),\ k = 2,3,4$ $\beta_p/\delta_p, \tau^2 \sim N(0, \delta_p \tau^2 + (1-\delta_p) u_0 \tau^2),$ $p = 1,5, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $w_{mk}/\gamma_{mk}, \beta_k, \sigma_k^2 \sim N(\gamma_{mk} \beta_k, \gamma_{mk} \sigma_k^2 + (1-\gamma_{mk}) u_0 \sigma_k^2),$ $k = 2,3,4,\ m = 1, \dots, 5$	$\gamma_{mk} / p_{mk} \sim \mathcal{B}(p_{mk}),$ $\Phi^{-1}(p_{mk}) = \varepsilon_{mk},$ $\boldsymbol{\varepsilon_k} / \sigma_k^2, \rho_k \sim \mathcal{N}_M(\mathbf{0}, \sigma_k^{-2}(\boldsymbol{R} - \rho_k \boldsymbol{\Omega})),$ $k = 2,3,4,\ m = 1, \dots, 5$	Model 7 with SSVS imposed on fixed effects
9	$\beta_p \sim \mathcal{N}(0, 0.1),\ p = 1, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $w_{mk}/\gamma_{mk}, \beta_k, \sigma_k^2 \sim N(\gamma_{mk} \beta_k, \gamma_{mk} \sigma_k^2 + (1-\gamma_{mk}) u_0 \sigma_k^2),$ $w_{mk} = F^{-1}(\Phi(\theta_{mk})),\ \theta_k / \rho_k \sim \mathcal{N}_M(\mathbf{0}, \boldsymbol{R} - \rho_k \boldsymbol{\Omega}),\ k = 2,3,4,\ m = 1, \dots, 5$	$\gamma_{mk} / p_k \sim \mathcal{B}(p_k),\ p_k \sim \mathcal{U}(0,1),\ k = 2,3,4,\ m = 1, \dots, 5$	Local BVS method utilizing the a Gaussian copula
10	$\beta_k \sim \mathcal{N}(0, 0.1),\ k = 2,3,4$ $\beta_p/\delta_p, \tau^2 \sim N(0, \delta_p \tau^2 + (1-\delta_p) u_0 \tau^2),$ $p = 1,5, \dots, 10$	$\boldsymbol{w_0} \sim \mathcal{N}_N(\mathbf{0}, (\sigma_0^2 \exp(-\varphi \boldsymbol{\Lambda}))^{-1})$ $w_{mk}/\gamma_{mk}, \beta_k, \sigma_k^2 \sim N(\gamma_{mk} \beta_k, \gamma_{mk} \sigma_k^2 + (1-\gamma_{mk}) u_0 \sigma_k^2),$ $w_{mk} = F^{-1}(\Phi(\theta_{mk})),\ \theta_k / \rho_k \sim \mathcal{N}_M(\mathbf{0}, \boldsymbol{R} - \rho_k \boldsymbol{\Omega}),\ k = 2,3,4,\ m = 1, \dots, 5$	$\gamma_{mk} / p_k \sim \mathcal{B}(p_k),\ p_k \sim \mathcal{U}(0,1),\ k = 2,3,4,\ m = 1, \dots, 5$	Model 9 with SSVS imposed on fixed effects

Table 6.2: Proportional inclusion of the fix components for models 1-10, based on SSVS. Results were derived from 20 simulations.

Fixed effects	Model 2	Model 4	Model 6	Model 8	Model 10
β_1	100%	100%	100%	100%	100%
β_2	100%	*	*	*	*
β_3	100%	*	*	*	*
β_4	100%	*	*	*	*
β_5	100%	100%	100%	100%	100%
β_6	100%	100%	100%	100%	100%
β_7	0%	0%	0%	0%	0%
β_8	0%	5%	5%	10%	10%
β_9	0%	0%	0%	0%	0%
β_{10}	5%	10%	10%	10%	10%

* Variables not included in SSVS.

Table 6.3: Proportional posterior inclusions for the spatially varying coefficients β_2, β_3 and β_4 and their local effects, stratified by variable selection. Results were derived from 20 simulations.

Indicator	$\beta_2\ (k=2)$						$\beta_3\ (k=3)$						$\beta_4\ (k=4)$					
	M*5	M6	M7	M8	M9	M10	M5	M6	M7	M8	M9	M10	M5	M6	M7	M8	M9	M10
$\gamma_k = \binom{0}{0}$	0%	0%					0%	0%					0%	0%				
$\gamma_k = \binom{1}{0}$	30%	35%					55%	60%					40%	50%				
$\gamma_k = \binom{1}{1}$	70%	65%					45%	40%					60%	50%				
γ_{1k}			50%	45%	25%	30%			50%	45%	30%	30%			30%	30%	25%	25%
γ_{2k}			35%	35%	25%	20%			35%	35%	25%	20%			15%	20%	10%	15%
γ_{3k}			60%	60%	40%	45%			55%	60%	35%	40%			0%	0%	5%	0%
γ_{4k}			60%	60%	30%	35%			60%	65%	40%	30%			0%	0%	10%	10%
γ_{5k}			30%	20%	15%	10%			30%	25%	10%	10%			45%	45%	45%	45%

* M stands for model.

Table 6.4: Description of child, maternal and household interventions coverage indicators utilized in the application.

Interventions	Definition
Use of ITNs from under-five children	Proportion of under the age of five children who slept under an ITN the night before the survey
Antimalarial medication	Proportion of under the age of five children with fever who received any antimalarial medication
Breastfeeding within 24 hours	Proportion of new-born children which was breastfed within one day after birth
Postnatal check	Proportion of postnatal check-up for infants within two months after birth
Measles immunization	Proportion of children vaccinated against measles
DPT3 immunization	Proportion of children which was vaccinated against measles
Vitamin A supplementation	Proportion of children which was vaccinated with DPT3
Improved sanitation	Proportion of household using improved sanitation facilities

Table 6.5: Posterior mean inclusion probabilities for each variable and model applied on the Burundi DHS data. Asterisk (*) indicates no SSVS for the corresponding covariates.

Variable (indicator)	Model 2	Model 4	Model 6	Model 8	Model 10
Sex (δ_1)	0.30	0.31	0.31	0.29	0.30
Use of ITNs (δ_2)	**0.74**	*	*	*	*
Antimalarial (δ_3)	**0.80**	*	*	*	*
Breastfeeding (δ_4)	**0.59**	*	*	*	*
Postnatal check (δ_5)	0.37	0.37	0.36	0.36	0.37
Measles (δ_6)	**0.51**	**0.53**	**0.55**	0.46	**0.53**
DPT3 (δ_7)	**1**	**1**	**1**	**1**	**1**
Vitamin A (δ_8)	**1**	**1**	**1**	**1**	**1**
Improved Sanitation (δ_9)	**0.88**	**0.66**	**0.65**	**0.61**	**0.87**
Rural-urban (δ_{10})	0.12	0.14	0.13	0.13	0.12
Delivery place (δ_{11})	0.49	0.41	0.44	0.42	0.45
Birth order (δ_{12})	**0.78**	**0.79**	**0.82**	**0.85**	**0.80**
Mothers age (δ_{13})	0.32	0.34	0.34	0.34	0.34
Mothers live births (δ_{14})	**0.88**	**0.90**	**0.92**	**0.93**	**0.91**
Mothers educ. level (δ_{15})	0.08	0.09	0.09	0.09	0.09
Asset Index (δ_{16})	**0.96**	**0.97**	**0.97**	**0.96**	**0.96**
Altitude (δ_{17})	0.09	0.09	0.09	0.08	0.08
NDVI (δ_{18})	0.08	0.08	0.09	0.08	0.07
LSTN (δ_{19})	0.11	0.09	0.11	0.09	0.08
LSTD (δ_{20})	0.33	0.22	0.28	0.21	0.17
Rainfall (δ_{21})	0.05	0.06	0.06	0.06	0.06
Forests (δ_{22})	0.42	0.43	0.43	0.41	0.42
Grasslands (δ_{23})	0.44	0.44	0.44	0.44	0.43
Croplands (δ_{24})	0.16	0.22	0.22	0.20	0.16

Table 6.6: Posterior mean inclusion probabilities for the spatially varying coefficients β_2, β_3 and β_4 and their local effects, stratified by variable selection. Results were derived from the initial fit to Burundi DHS 2016.

Indicator	β_2 $(k=2)$						β_3 $(k=3)$						β_4 $(k=4)$					
	M*5	M6	M7	M8	M9	M10	M5	M6	M7	M8	M9	M10	M5	M6	M7	M8	M9	M10
$\gamma_k = \binom{0}{0}$	0	0					1	0					0	0				
$\gamma_k = \binom{1}{0}$	1	1					0	1					1	0				
$\gamma_k = \binom{1}{1}$	0	0					0	0					0	1				
Bubanza (γ_{1k})			0.46	**0.52**	0.12	0.07			0.42	0.38	0.14	0.09			0.48	0.41	0.05	0.05
Bujumbura Mairie (γ_{2k})			0.43	**0.53**	0.07	0.05			0.45	0.41	0.12	0.07			0.48	0.40	0.04	0.05
Bujumbura Rural (γ_{3k})			0.44	**0.52**	0.11	0.07			0.42	0.46	0.16	0.09			0.46	0.42	0.04	0.06
Bururi (γ_{4k})			0.42	**0.56**	0.10	0.05			0.42	0.46	0.13	0.08			0.49	0.47	0.06	0.05
Cankuzo (γ_{5k})			0.40	**0.53**	0.09	0.05			0.39	0.40	0.13	0.10			0.37	0.38	0.04	0.03
Cibitoke (γ_{6k})			0.34	**0.52**	0.08	0.04			0.39	0.39	0.10	0.07			0.34	0.33	0.02	0.02
Gitega (γ_{7k})			0.46	**0.60**	0.12	0.16			**0.53**	**0.51**	**0.56**	**0.64**			0.41	0.38	0.16	0.12
Karuzi (γ_{8k})			0.37	**0.51**	0.08	0.06			0.31	0.40	0.11	0.06			0.37	0.32	0.03	0.05
Kayanza (γ_{9k})			0.35	**0.50**	0.08	0.06			0.41	0.37	0.13	0.07			0.35	0.36	0.03	0.05
Kirundo (γ_{10k})			0.39	0.48	0.09	0.06			0.40	0.44	0.12	0.10			0.36	0.37	0.03	0.03
Makamba (γ_{11k})			0.38	**0.51**	0.08	0.04			0.39	0.43	0.09	0.09			0.45	0.43	0.05	0.05
Muramvya (γ_{12k})			0.41	**0.54**	0.14	0.12			0.47	0.43	0.26	0.19			0.45	0.37	0.05	0.07
Muyinga (γ_{13k})			0.34	0.48	0.08	0.03			0.34	0.40	0.12	0.07			0.35	0.32	0.02	0.02
Mwaro (γ_{14k})			0.43	**0.51**	0.13	0.08			0.46	0.44	0.16	0.10			0.42	0.46	0.09	0.08
Ngozi (γ_{15k})			0.35	0.48	0.08	0.05			0.36	0.38	0.13	0.09			0.39	0.37	0.04	0.04
Rutana (γ_{16k})			0.44	**0.51**	0.11	0.06			0.41	0.42	0.15	0.08			0.47	0.41	0.04	0.09
Ruyigi (γ_{17k})			0.45	**0.58**	0.25	0.15			0.46	0.39	0.22	0.13			0.48	0.42	0.12	0.13

* M stands for model

Table 6.7: Posterior summaries of regression coefficients. Results are derived from the re-fit on the Burundi DHS 2016 data, based on the variables selected from the SSVS.

	Model A	Model B	Model C	Model D	Model E	Model F	Model G
Use of ITNs	-0.98 (-2.09,0.10)	-0.62 (-1.15,0.02)	**-1.67 (-2.67,-0.63)**	-0.83 (-3.14,1.18)	-1.49 (-4.10,0.57)	**-0.66 (-1.58,-0.03)**	-1.38 (-3.33,0.50)
Antimalarial	**0.67 (0.16,1.14)**	0.20 (-0.86,0.79)	Not selected[1]	0.69 (-0.10,1.54)	-0.16 (-1.03,0.54)	-0.37 (-1.08,0.30)	0.05 (-0.95,0.66)
Breastfeeding	-1.49 (-3.04,0.24)	-0.42 (-0.87,0.08)	-1.09 (-2.88,0.16)	-0.49 (-1.46,0.48)	-1.25 (-3.99,0.23)	-1.18 (-3.01,0.24)	-0.89 (-2.22,0.41)
Measles	-0.28 (-1.48,1.05)	-0.09 (-1.23,0.97)	-0.31 (-1.48,0.93)	0.03 (-1.23,1.27)	Not selected[2]	Not selected[2]	-0.03 (-1.29,1.22)
DPT3	**-3.13 (-4.57,-1.93)**	**-3.15 (-4.27,-1.89)**	**-3.07 (-4.29,-1.98)**	**-3.40 (-4.64,-1.99)**	**-3.28 (-4.19,-2.49)**	**-3.18 (-4.04,-2.34)**	**-3.21 (-4.57,-1.93)**
Vitamin A	**-0.93 (-1.33,-0.53)**	**-1.15 (-1.58,-0.65)**	**-1.01 (-1.42,-0.62)**	**-1.09 (-1.55,-0.62)**	**-1.09 (-1.55,-0.68)**	**-1.19 (-1.63,-0.75)**	**-1.11 (-1.60,-0.63)**
Improved sanitation	**-0.49 (-0.78,-0.20)**	**-0.41 (-0.80,-0.04)**	**-0.59 (-0.87,-0.30)**	**-0.42 (-0.77,-0.06)**	**-0.46 (-0.81,-0.08)**	**-0.43 (-0.81,-0.07)**	**-0.46 (-0.83,-0.08)**
Birth order	**-0.46 (-0.76,-0.14)**	**-0.46 (-0.75,-0.14)**	**-0.45 (-0.74,-0.15)**	**-0.46 (-0.77,-0.14)**	**-0.46 (-0.76,-0.15)**	**-0.46 (-0.75,-0.16)**	**-0.46 (-0.76,-0.15)**
Mothers live births	**0.56 (0.27,0.83)**	**0.57 (0.29,0.84)**	**0.55 (0.26,0.82)**	**0.57 (0.29,0.85)**	**0.57 (0.29,0.84)**	**0.57 (0.29,0.84)**	**0.58 (0.29,0.85)**
Asset Index	**0.29 (0.13,0.45)**	**0.31 (0.15,0.46)**	**0.32 (0.16,0.47)**	**0.30 (0.14,0.46)**	**0.31 (0.15,0.47)**	**0.31 (0.15,0.47)**	**0.30 (0.15,0.46)**
WAIC	8619.5	8624.6	8623.7	8616.0	8618.2	8620.1	8619.1
Parent models	1 – 2	3 – 4	5	6	7	8	9 – 10

1. Not selected based on the Global Bayesian Variable selection method; 2. Not selected based on the Stochastic Search Variable Selection method.

6.6 Appendix

Table 6.8: Pre-determined parameter values and distributions used to generate our simulation (part I).

	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=7$	$p=8$	$p=9$	$p=10$
β_p	-0.80	0.80	-1	-1.10	0.50	1	0.70	0	0	0	0
x_{ip}	1	$\mathcal{N}(0,0.50)$	$\mathcal{N}(0,0.50)$	$\mathcal{N}(0,0.50)$	$\mathcal{B}(0.5)$	$\mathcal{B}(0.5)$	$\mathcal{N}(1,0.50)$	$\mathcal{N}(-1,2)$	$\mathcal{B}(0.5)$	$\mathcal{N}(1,2)$	$\mathcal{B}(0.5)$

$\mathcal{N}$ stands for Normal distribution; $\mathcal{B}$ stands for Bernoulli distribution.

Table 6.9: Pre-determined parameter values used to generate our simulation (part II).

	α	σ_0^2	φ	σ_1^{-2}	ρ_1	σ_2^{-2}	ρ_2	σ_3^{-2}	ρ_3
Interpretation	Shape parameter	Variance of GP	Spatial decay	Variance of CAR	Spatial association	Variance of CAR	Spatial association	Variance of CAR	Spatial association
Value	1	1	0.20	2	0.50	2.5	0.70	2	0.60

GP stands for Gaussian process.

Table 6.10: Initial values used for the two chains of the MCMC.

	Chain 1	Chain 2		Chain 1	Chain 2
β_1	0.80	1	β_0	-0.80	0.50
β_2	-1	1	α	1	1
β_3	-1.10	1	σ_0^2	1	0.33
β_4	0.50	1	φ	0.20	5
β_5	1	1	σ_2^{-2}	2.5	1
β_6	0.70	1	ρ_2	0.70	0.90
β_7	0	1	σ_3^{-2}	2	1.5
β_8	0	1	ρ_3	0.60	0.40
β_9	0	1	σ_4^{-2}	1	3
β_{10}	0	1	ρ_4	0.70	0.80

Table 6.11: Proportional inclusion of the fix components for models 1-10, based on SSVS. Results were derived from 20 simulations.

Fixed effects	Model 2	Model 4	Model 6	Model 8	Model 10
β_1	20/20	20/20	20/20	20/20	20/20
β_2	20/20	*	*	*	*
β_3	20/20	*	*	*	*
β_4	20/20	*	*	*	*
β_5	20/20	20/20	20/20	20/20	20/20
β_6	20/20	20/20	20/20	20/20	20/20
β_7	0/20	0/20	0/20	0/20	0/20
β_8	0/20	1/20	1/20	2/20	2/20
β_9	0/20	0/20	0/20	0/20	0/20
β_{10}	1/20	2/20	2/20	2/20	2/20

* Variables not included in SSVS.

Table 6.12: Proportional posterior inclusion for the spatially varying coefficients $\boldsymbol{\beta_2}$, $\boldsymbol{\beta_3}$ and $\boldsymbol{\beta_4}$ and their local effects, stratified by variable selection. Results were derived from 20 simulations.

	β_2						β_3						β_4					
Indicator	M*5	M6	M7	M8	M9	M10	M5	M6	M7	M8	M9	M10	M5	M6	M7	M8	M9	M10
$\gamma_k = \binom{0}{0}$	0/20	0/20					0/20	0/20					0/20	0/20				
$\gamma_k = \binom{1}{0}$	6/20	7/20					11/20	12/20					8/20	10/20				
$\gamma_k = \binom{1}{1}$	14/20	13/20					9/20	8/20					12/20	10/20				
γ_{1k}			10/20	9/20	5/20	6/20			10/20	9/20	6/20	6/20			6/20	6/20	5/20	5/20
γ_{2k}			7/20	7/20	5/20	4/20			7/20	7/20	5/20	4/20			3/20	4/20	2/20	3/20
γ_{3k}			12/20	12/20	8/20	9/20			11/20	12/20	7/20	8/20			0/20	0/20	1/20	0/20
γ_{4k}			12/20	12/20	6/20	7/20			12/20	13/20	8/20	6/20			0/20	0/20	2/20	2/20
γ_{5k}			6/20	4/20	3/20	2/20			6/20	5/20	2/20	2/20			9/20	9/20	9/20	9/20

* M stand for model

Table 6.13: Summary of indicators derived from the local BVS methods with posterior mean inclusion probability greater than 50%.

Sim*	Model 3	Model 4	Model 7	Model 8	Model 9	Model 10
1	γ_{32}** ,γ_{42},γ_{33},γ_{43}	γ_{32},γ_{42},γ_{33},γ_{43}	γ_{32},γ_{42},γ_{33},γ_{43}, γ_{54}	γ_{32},γ_{42},γ_{33},γ_{43}, γ_{54}	γ_{32},γ_{42},γ_{33},γ_{43}, γ_{54}	γ_{32},γ_{42},γ_{33},γ_{43}, γ_{54}
2	γ_{23}, γ_{33}, γ_{43}	γ_{22}, γ_{23}	-	-	-	-
3	γ_{43}	-	γ_{12},γ_{42},γ_{13},γ_{43}, γ_{54}	γ_{12},γ_{42},γ_{13},γ_{43}, γ_{54}	γ_{12},γ_{42},γ_{13},γ_{54}	γ_{12},γ_{42},γ_{13},γ_{54}
4	γ_{22},γ_{42},γ_{23}	γ_{22},γ_{42},γ_{23}	γ_{22},γ_{42},γ_{52},w_{23}, γ_{43}, γ_{53},γ_{14}	γ_{22},γ_{42},γ_{23}, γ_{43},γ_{53},γ_{14}	γ_{14}	γ_{14}
5	-	-	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{33},γ_{43},γ_{24},γ_{54}	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{33},γ_{43},γ_{24},γ_{54}	γ_{12},γ_{32},γ_{42},γ_{52}, γ_{13},γ_{33},γ_{43}, γ_{24},γ_{34},γ_{44},γ_{54}	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{33},γ_{43},γ_{24},γ_{54}
6	-	-	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{33},γ_{43},γ_{54}	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{33},γ_{43},γ_{54}	γ_{54}	γ_{54}
7	γ_{43}	γ_{43}	γ_{52},γ_{53},γ_{24}	γ_{43},γ_{24}	-	-
8	γ_{22},γ_{23}	γ_{22}	γ_{12},γ_{22},γ_{13},γ_{23}, γ_{54}	γ_{12},γ_{22},γ_{13},γ_{23}, γ_{54}	γ_{12},γ_{22},γ_{13},γ_{23}, γ_{54}	γ_{12},γ_{22},γ_{13},γ_{23}, γ_{54}
9	γ_{42},γ_{43}	γ_{42},γ_{43}	γ_{32},γ_{42},γ_{13},γ_{33}, γ_{43},γ_{24}	γ_{32},γ_{42},γ_{33},γ_{43}, γ_{24}	γ_{42},γ_{43},γ_{24}	γ_{42},γ_{43},γ_{24}
10	γ_{23},γ_{43}	γ_{22},γ_{42},γ_{23},γ_{33}, γ_{43}	γ_{12},γ_{22},γ_{32},γ_{42}, γ_{13},γ_{23},γ_{43},γ_{54}	γ_{12},γ_{22},γ_{32},γ_{42}, γ_{13},γ_{23},γ_{33},γ_{43}, γ_{54}	γ_{54}	γ_{54}
11	γ_{12},γ_{32},γ_{13},γ_{23}, γ_{33}	γ_{12},γ_{13},γ_{23},γ_{33}	γ_{12},γ_{22},γ_{32}, γ_{13},γ_{23},γ_{33},γ_{54}	γ_{12},γ_{22},γ_{32}, γ_{13},γ_{23},γ_{33},γ_{54}	γ_{12},γ_{22},γ_{32},γ_{13}, γ_{23},γ_{33},γ_{54}	γ_{12},γ_{22},γ_{32},γ_{13}, γ_{23},γ_{33},γ_{44},γ_{54}
12	γ_{42}	γ_{42}	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{33},γ_{43},γ_{54}	γ_{12},γ_{32},γ_{42}, γ_{13},γ_{33},γ_{43},γ_{54}	γ_{42},γ_{13},γ_{43},γ_{54}	γ_{12},γ_{32},γ_{42},γ_{13}, γ_{43},γ_{54}
13	γ_{32},γ_{33},γ_{43}	γ_{32},γ_{33}	γ_{32},γ_{42},γ_{52}, γ_{33},γ_{43},γ_{53},γ_{14}	γ_{32},γ_{42},γ_{52},γ_{33}, γ_{43},w_{53},γ_{14}	γ_{32},γ_{33},γ_{43},γ_{14}	γ_{32},γ_{33},γ_{14}
14	γ_{52},γ_{53}	γ_{22},γ_{52},γ_{23},γ_{53}	γ_{52},γ_{53},γ_{14}	γ_{52},γ_{53},γ_{14}	γ_{52},γ_{53}	γ_{52},γ_{53}
15	γ_{23}	γ_{23}	γ_{12},γ_{22},γ_{13},γ_{23}	γ_{12},γ_{22},γ_{13},γ_{23}	γ_{22},γ_{23}	γ_{22},γ_{23}
16	γ_{13}	γ_{13}	γ_{12},γ_{32},γ_{13},γ_{33}, γ_{54}	γ_{12},γ_{32},γ_{13},γ_{33}, γ_{54}	γ_{12},γ_{32},γ_{52},γ_{13}, γ_{23},γ_{33},γ_{14},γ_{44}, γ_{54}	γ_{12},γ_{32},γ_{13},γ_{33}, γ_{14},γ_{44},γ_{54}
17	γ_{32},γ_{33}	γ_{22},γ_{32},γ_{23},γ_{33}	γ_{22},γ_{32},γ_{52},γ_{23}, γ_{33},γ_{53},γ_{14}	γ_{22},γ_{32},γ_{52}, γ_{23},γ_{33},γ_{53},γ_{14}	γ_{32},γ_{33},γ_{14}	γ_{32},γ_{33},γ_{14}
18	γ_{42},γ_{33},γ_{43}	γ_{42},γ_{43}	γ_{32},γ_{42},γ_{52}, γ_{33},γ_{43},γ_{53},γ_{14}	γ_{32},γ_{42},γ_{52}, γ_{33},γ_{43},γ_{53},γ_{14}, γ_{24}	γ_{32},γ_{42},γ_{52}, γ_{43},γ_{53}	γ_{32},γ_{42},γ_{52},γ_{33}, γ_{43},γ_{53},γ_{24}
19	γ_{42},γ_{33},γ_{43}	γ_{42},γ_{43}	γ_{22},γ_{32},γ_{42},γ_{23}, γ_{33},γ_{43},γ_{14}	γ_{22},γ_{32},γ_{42},γ_{23}, γ_{33},γ_{43},γ_{14}	γ_{22},γ_{32},γ_{42},γ_{23}, γ_{33},γ_{43},γ_{14}	γ_{22},γ_{32},γ_{42},γ_{23}, γ_{33},γ_{43},γ_{14}
20	γ_{22},γ_{42},γ_{43}	γ_{22},γ_{42},γ_{43}	γ_{42},γ_{43}	γ_{42},γ_{43}	-	-

* Simulation

** γ_{mk}, $m = 1, \ldots, 5$ represents the province and $k = 2,3,4$ the spatially varying coefficient

Table 6.14: Posterior summaries of regression coefficients and the shape parameter alpha of Weibull lifetime distribution. Results are derived from the initial fit to Burundi DHS 2016.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Shape (α)	0.33 (0.31,0.36)	0.34 (0.31,0.36)	0.34 (0.31,0.36)	0.33 (0.31,0.36)	0.33 (0.31,0.36)	0.34 (0.31,0.36)	0.33 (0.31,0.36)	0.33 (0.31,0.36)	0.33 (0.31,0.36)	0.34 (0.31,0.36)
Sex (β_1)	**0.15 (0,0.3)**	0.04 (-0.03,0.25)	**0.15 (0.01,0.3)**	0.04 (-0.03,0.25)	**0.15 (0,0.3)**	0.04 (-0.03,0.26)	**0.15 (0,0.3)**	0.04 (-0.03,0.25)	**0.15 (0,0.3)**	0.04 (-0.03,0.25)
Use of ITNs (β_2)	-0.88 (-2.2,0.39)	-0.72 (-2.07,0.06)	-0.29 (-2.09,1.11)	**-1.72 (-2.34,-0.81)**	-0.62 (-1.64,1.04)	-0.57 (-1.23,1.33)	-0.21 (-1.56,0.66)	-0.4 (-3.98,0.59)	-0.61 (-2.78,0.36)	-0.61 (-2.07,0.67)
Antimalarial (β_3)	0.41 (-0.2,0.98)	0.57 (-0.03,1.14)	-0.49 (-2.23,1.53)	-0.75 (-1.55,1.48)	0 (-0.41,0.55)	-0.4 (-0.97,1.48)	0.12 (-2.17,1.02)	0.12 (-1.66,1.03)	0.41 (-0.86,1.14)	0.51 (-0.14,1.21)
Breastf. (β_4)	**-1.64 (-2.91,-0.14)**	-0.22 (-2.37,0.15)	-0.65 (-1.46,0.13)	-0.09 (-1.32,0.77)	-0.65 (-1.6,0.1)	-0.35 (-1.49,0.36)	-0.39 (-2.15,0.4)	-0.38 (-2.08,0.37)	-1.04 (-2.78,0.38)	-0.8 (-2.08,0.43)
Postnatal check (β_5)	-0.48 (-1.63,0.64)	-0.01 (-1.05,0.45)	0.12 (-1.25,1.46)	0 (-0.8,0.88)	-0.2 (-1.42,1.01)	0 (-0.84,0.76)	0.07 (-1.29,1.42)	0 (-0.82,0.84)	-0.21 (-1.54,1.23)	0 (-0.93,0.69)
Measles (β_6)	-0.06 (-1.29,1.3)	-0.05 (-1.39,0.24)	-0.01 (-1.3,1.28)	-0.05 (-1.39,0.24)	0 (-1.23,1.23)	-0.06 (-1.46,0.19)	0.05 (-1.13,1.3)	-0.03 (-1.28,0.29)	0.02 (-1.18,1.24)	-0.05 (-1.39,0.34)
DPT3 (β_7)	**-3.37 (-4.71,-2.1)**	**-2.76 (-3.92,-1.54)**	**-3.17 (-4.52,-1.87)**	**-2.51 (-3.57,-1.26)**	**-3.24 (-4.68,-2.1)**	**-2.52 (-3.59,-1.34)**	**-3.34 (-4.65,-2.06)**	**-2.62 (-3.68,-1.34)**	**-3.36 (-4.6,-2.1)**	**-2.61 (-3.73,-1.49)**
Vitamin A (β_8)	**-0.95 (-1.38,-0.5)**	**-0.95 (-1.36,-0.55)**	**-1.12 (-1.6,-0.64)**	**-1.11 (-1.56,-0.65)**	**-1.03 (-1.5,-0.58)**	**-1.07 (-1.52,-0.64)**	**-1.18 (-1.66,-0.68)**	**-1.16 (-1.62,-0.69)**	**-1.09 (-1.56,-0.61)**	**-1.02 (-1.46,-0.58)**
Improved Sanit. (β_9)	**-0.47 (-0.8,-0.15)**	-0.41 (-0.74,0.01)	**-0.39 (-0.78,0)**	-0.28 (-0.77,0.04)	**-0.45 (-0.8,-0.1)**	-0.27 (-0.73,0.04)	-0.37 (-0.76,0.02)	-0.21 (-0.71,0.04)	**-0.43 (-0.77,-0.07)**	-0.42 (-0.77,0.01)
Rural-urban (β_{10})	0.02 (-0.24,0.29)	0 (-0.11,0.12)	0.07 (-0.2,0.36)	0.01 (-0.08,0.2)	0.04 (-0.23,0.31)	0 (-0.09,0.16)	0.07 (-0.2,0.37)	0 (-0.09,0.17)	0.05 (-0.23,0.33)	0 (-0.11,0.12)
Delivery place (β_{11})	**0.23 (0.03,0.43)**	0.07 (-0.03,0.4)	**0.21 (0.01,0.41)**	0.05 (-0.04,0.37)	**0.23 (0.01,0.42)**	0.05 (-0.04,0.38)	**0.22 (0.01,0.42)**	0.05 (-0.04,0.37)	**0.22 (0.01,0.43)**	0.06 (-0.03,0.38)
Birth order (β_{12})	**-0.46 (-0.77,-0.15)**	-0.38 (-0.71,0.04)	**-0.47 (-0.78,-0.16)**	-0.38 (-0.72,0.03)	**-0.47 (-0.77,-0.16)**	-0.41 (-0.73,0.03)	**-0.48 (-0.78,-0.16)**	-0.42 (-0.73,0.03)	**-0.47 (-0.77,-0.16)**	-0.38 (-0.72,0.03)
Mothers age (β_{13})	0.32 (-0.29,0.83)	0.02 (-0.15,0.64)	0.33 (-0.28,0.86)	0.02 (-0.14,0.67)	0.33 (-0.28,0.84)	0.02 (-0.12,0.67)	0.34 (-0.28,0.86)	0.02 (-0.13,0.67)	0.34 (-0.27,0.86)	0.02 (-0.14,0.67)
Mothers live births (β_{14})	**0.56 (0.27,0.83)**	**0.48 (0,0.79)**	**0.58 (0.28,0.85)**	**0.5 (0.01,0.8)**	**0.57 (0.28,0.84)**	**0.51 (0.01,0.8)**	**0.58 (0.29,0.85)**	**0.53 (0.02,0.82)**	**0.57 (0.28,0.83)**	**0.5 (0.01,0.8)**
Mothers educ. level (β_{15})	0.03 (-0.13,0.18)	0 (-0.06,0.08)	0.03 (-0.13,0.19)	0.01 (-0.06,0.1)	0.03 (-0.13,0.19)	0 (-0.06,0.08)	0.03 (-0.13,0.19)	0 (-0.06,0.09)	0.03 (-0.13,0.19)	0 (-0.06,0.1)
Asset Index (β_{16})	**0.28 (0.12,0.45)**	**0.3 (0.07,0.47)**	**0.29 (0.12,0.46)**	**0.31 (0.09,0.47)**	**0.3 (0.13,0.46)**	**0.31 (0.09,0.47)**	**0.29 (0.12,0.45)**	**0.31 (0.07,0.47)**	**0.28 (0.12,0.45)**	**0.3 (0.06,0.46)**
Altitude (β_{17})	-0.01 (-0.25,0.21)	0.01 (-0.06,0.09)	-0.08 (-0.33,0.17)	0 (-0.08,0.07)	-0.01 (-0.27,0.23)	0 (-0.07,0.08)	-0.09 (-0.34,0.17)	0 (-0.08,0.07)	-0.04 (-0.29,0.21)	0 (-0.08,0.06)

NDVI (β_{18})	-0.04 (-0.15,0.08)	-0.02 (-0.09,0.04)	-0.04 (-0.16,0.08)	-0.01 (-0.09,0.04)	-0.01 (-0.13,0.11)	-0.01 (-0.09,0.04)	-0.03 (-0.16,0.1)	-0.02 (-0.09,0.04)	-0.03 (-0.15,0.09)	-0.01 (-0.08,0.04)
LSTN (β_{19})	-0.1 (-0.33,0.12)	-0.01 (-0.15,0.05)	-0.16 (-0.39,0.09)	0 (-0.11,0.06)	-0.16 (-0.4,0.08)	-0.01 (-0.14,0.05)	-0.17 (-0.41,0.07)	-0.01 (-0.11,0.06)	-0.13 (-0.37,0.12)	0 (-0.07,0.07)
LSTD (β_{20})	0.14 (-0.03,0.32)	0.05 (-0.02,0.24)	0.14 (-0.05,0.32)	0.03 (-0.03,0.21)	**0.21 (0.03,0.38)**	0.03 (-0.03,0.22)	0.15 (-0.04,0.33)	0.03 (-0.03,0.2)	0.14 (-0.05,0.33)	0.03 (-0.03,0.16)
Rainfall (β_{21})	-0.02 (-0.09,0.05)	-0.01 (-0.06,0.04)	-0.02 (-0.1,0.07)	0 (-0.06,0.05)	-0.02 (-0.1,0.06)	-0.01 (-0.06,0.05)	-0.02 (-0.1,0.07)	-0.01 (-0.07,0.04)	-0.02 (-0.1,0.06)	-0.01 (-0.07,0.04)
Forests (β_{22})	0.68 (-1.47,2.55)	0.01 (-0.94,1.3)	0.33 (-1.84,2.31)	0 (-1.12,1.09)	0.61 (-1.57,2.54)	0 (-1.08,1.08)	0.29 (-1.95,2.26)	0 (-1.17,1)	0.58 (-1.62,2.46)	0 (-0.94,1.24)
Grasslands (β_{23})	-0.45 (-2.71,1.66)	-0.01 (-1.54,0.73)	-0.59 (-3,1.74)	-0.01 (-1.46,0.74)	-0.26 (-2.61,1.94)	-0.01 (-1.45,0.72)	-0.51 (-2.98,1.77)	-0.01 (-1.53,0.74)	-0.47 (-2.84,1.73)	-0.01 (-1.38,0.8)
Croplands (β_{24})	-0.04 (-0.4,0.34)	0 (-0.23,0.17)	-0.21 (-0.69,0.28)	-0.01 (-0.43,0.12)	-0.1 (-0.53,0.31)	-0.01 (-0.38,0.18)	-0.23 (-0.72,0.25)	-0.01 (-0.39,0.13)	-0.09 (-0.57,0.33)	0 (-0.21,0.17)
Constant term (β_0)	-0.07 (-1.75,1.36)	-1.52 (-2.69,0.36)	0.35 (-4.89,4.86)	1.02 (-2.66,4.78)	-0.54 (-3.52,2.97)	0.08 (-5.68,3)	-1.34 (-2.58,4.79)	-1.26 (-2.83,3.99)	-0.6 (-2.31,1.22)	-1.18 (-3.03,0.12)

Chapter 7

Discussion

The present PhD thesis has been committed to the spatio-temporal modelling of under-five mortality data in the African setting using primarily routinely collected, cross-sectional, household-based survey data coming from the Demographic and Health surveys program. In the absence of improved vital registration systems and strong health infrastructure, as well as having the world's highest under-five mortality burden, our aim was to contribute to the scientific fields of Epidemiology and Statistics by modelling and analysing preschool mortality data for most of the African countries. A list of key achievements succeeded by this PhD thesis are as follows:

(1) We proposed a novel indicator, namely the malaria-anemia comorbidity prevalence indicator, in its capacity to estimate malaria-related mortality for under the age of five children in sub-Saharan Africa.

(2) We obtained maps at high spatial resolution of the malaria-anemia comorbidity effect on all-cause under-five mortality.

(3) We assessed the sensitivity of Bayesian variable selection methods for geostatistical models with spatially varying coefficients.

(4) We associated important preventative or curative health interventions with all-cause under-five mortality and identified health inequities between and within countries in the African continent, based on the geographical difference of intervention effects.

(5) We quantified the geographical distribution of the contribution of the leading causes of under-five mortality in Africa with febrile response.

(6) We developed a novel model formulation to estimate the impact of health interventions on the temporal changes in under-five mortality risk for 21 African countries.

Detailed information on the backgrounds, methods, analysis and discussions of the aforementioned achievements can be found in Chapters 2–6. The role of this section is to provide an overview of the main outcomes of our work, assess its limitations and provide recommendations for future research.

7.1 Major contributions of the work

7.1.1 Epidemiology

One of the most significant outcomes of the present PhD thesis is the proposed malaria-anemia comorbidity prevalence indicator as a novel measure to quantify malaria-related mortality in Africa. A global renewed interest on malaria control and elimination essentially started in 1998 when the Roll Back Malaria (RBM), currently RBM Partnership to end malaria, was launched with a goal to halve the malaria burden by 2010 and eliminate malaria by 2030. Roughly at the same time, the Mapping Malaria risk in Africa/Atlas du Risqué de la Malaria en Afrique (MARA/ARMA) project was initiated with a mandate to collect malaria infection prevalence data, entomological inoculation rates, case-incidence information and data on *An. Gambiae* and *An. Funestus* malaria vectors. Following the establishment of the RBM and MARA projects, several studies utilized different approaches and data sources in order to quantify and map malaria-related mortality in the African setting. All of these studies contributed to the scientific knowledge on malaria burden and provided key information to decision makers. Yet, given that most of these studies utilized survey data for their analysis, they faced a double challenge: firstly, finding a statistically important malaria-mortality association using the parasitaemia prevalence as an indicator to explain the malaria-mortality association and secondly, completely overlooked indirect causes of malaria mortality. The latter is a very important, as, for instance, one of the main outcomes of *Plasmodium* infection is anemia that leads to a serious life-threatening form of disease, i.e. severe malarial anemia, which is considered responsible for around one third of malaria deaths (Kai, 2008; Haldar, 2009). Additionally, malaria parasitaemia prevalence takes into account patients that experience, for example, uncomplicated malaria (presence of parasites without experiencing a life-threatening situation) and thus from a modelling perspective, many of the positives accounted in prevalence are not linked with mortality. Our work contributed to the development, proposal and validation of a novel indicator for quantifying malaria-mortality using survey data, which overcomes the above difficulties. Our main

conclusions were that malaria burden in sub-Saharan Africa is considerably underestimated when anemia in not taken into account and that the malaria-anemia comorbidity prevalence provides a useful measure of the malaria-related deaths. As the understanding of malaria-related, mortality provides insights in monitoring and evaluation of malaria control and elimination efforts, as well as in the planning of interventions, our work has the potential to impact key decisions made by policy makers. Further details are provided in Chapter 2. Another key result of the present PhD thesis is the modelling, mapping and presentation of health inequities at sub-national level in the African continent, due to variation in the association of preventative or curative health interventions with all-cause under-five mortality. Overall, Africa has seen a remarkable reduction in under-five deaths between 1990 and 2015, as the number of global deaths was nearly halved. Most of these gains were accomplished during the Millennium Development Goals era, especially towards the final half of that period when the scaling-up of preventative and curative health interventions took place. Modelling studies, field data and reports from local and international organisations have all pointed out that this particular increase in coverage of health interventions, accompanied by the profound effect of such interventions when properly implemented, constitute the reason of the unprecedented decline in pre-school mortality. However, it has recently attracted strong attention if there is an equal and fair distribution and effectiveness of such interventions across all areas covered by a country. For instance, a crucial question is if people leaving in remote or neglected areas of a given country experience health inequities compared to other subsets of the population residing in more well-off, urban areas. To this regard, several studies and national reports have described strong disparities in the under-five mortality rate, coverage of health interventions and some have raised the point of differences in the effectiveness of health interventions across various regions. In this thesis, we developed the first model-based estimates of sub-national health inequities due to the variation in the association of health interventions with all-cause under-five mortality over large scale in SSA. Based on the associations found, we provided a continental map highlighting these inequities at administrative level 1 for 28 countries in Africa. At the heart of

the Sustainable Development Goals is the reduction of inequalities and hence our work provides a roadmap on how to address these issues. Additionally, our work can inform decision makers about the areas of highest need, in terms of effectiveness of health interventions, and in doing so assist in malaria control and elimination efforts. Further details are provided in Chapter 3.

Results of our work inform about the contribution of the leading causes of under-five mortality in Africa on febrile response. One of the most common outcomes of inflammatory disease or infection is fever, albeit the identification of its cause is difficult. This is an important issue, as due to the large number of its potential causes, there are relatively high chances for wrong treatment of fever cases. Additionally, in the African setting, fever can be triggered by the leading causes of under-five mortality, i.e. acute respiratory infection, diarrhoea and malaria. Another crucial aspect of fever is that it is generally so strongly associated with malaria in endemic countries that often malaria is over-diagnosed and over-treated among under-five children with fever. Observational studies have reported that malaria can be the root-cause for less than one out of four children with fever. The current PhD thesis presents the first modelling study on the contribution of the leading causes of mortality on febrile response using routinely collected survey data for 16 sub-Saharan African countries. Our estimates confirm the strong contribution of diarrhoea and acute respiratory infection on fever and account only one out of five cases to malaria. These findings are of great public health significance, as they highlight the need for comprehensive testing before malaria is diagnosed and treated in endemic countries. Further details are provided in Chapter 4.

Chapter 5 provides information on the effect of 21 preventative or curative health interventions on changes in under-five mortality risk in 21 African countries. Since the remarkable reduction in the number of deaths for children below the age of five during the millennium Development Goals era, it is of major public health important to identify which interventions contributed the most on the decline of the under-five mortality risk. While several studies tried to estimate the effect of individual interventions on mortality, there is a

scarcity of studies taking a holistic approach and associating a large set of well-proven for their effectiveness health interventions with under-five mortality. Additionally, our study does not look at the effect of interventions in a cross-sectional fashion, but rather focuses on the difference in mortality risk between two distinct time points. Our results reflected the crucial role of vaccinations, micronutrient supplementation and treatments of diseases in the reduction of under-five mortality risk in the African setting.

Apart from the key take-home conclusions derived from this PhD thesis, as presented above, our work has contributed to the field of Epidemiology through several other aspects. At Chapter 2, we present high spatial resolution, i.e. 2x2 km^2, model-based risk map surfaces for the malaria-anemia comorbidity prevalence (both moderate and severe), malaria parasitaemia, at least moderate anemia and severe anemia for 16 sub-Saharan Africa countries. These estimates were derived using the latest DHS surveys for each country and hence are of great public health importance, in the context of identifying areas of severe disease. Chapter 3 contains an up-to-date, under-five mortality rate map of the African continent, presented at administrative level 1, again using data from the latest DHS surveys. Crucially, in that Chapter we present sets of continental and national maps with the spatially varying effect of 25 health interventions on all-cause, under-five mortality at sub-national level for 28 African countries. As with the disease maps, these are valuable cartographic efforts in terms of identifying subnational areas in which there is no statistical association between interventions and mortality. At chapter 4, we present continental maps with the geographical distribution of the association of acute respiratory infection, diarrhoea and malaria with fever. These maps show vividly the strong contribution of acute respiratory infection and diarrhoea to fever, while they highlight the varying effect of malaria to fever across different areas.

7.1.2 Statistics

This thesis contributed to the areas of statistics and data science through the development, implementation and interpretation of Bayesian geostatistical Weibull-survival and logistic-

regression models and their extensions to include spatially varying coefficients. The varying coefficient specification was implemented using three different specifications. In Chapters 3–6 we employed either an exchangeable or a conditional autoregressive structure for the province-level (Administrative level 1) effect of the varying coefficients. The exchangeable structure assumed that the province-level effects deviate from their national average independently, while the conditional autoregressive Gaussian process for the varying coefficients introduced spatial dependence among provinces based on their neighbouring structure. An interesting feature of our work is presented at Chapter 2, in which we fitted a varying coefficient model using a spatially continuous Gaussian process, on the malaria-anemia comorbidity effect, with Matérn covariance function. This specification allowed us to estimate the spatially varying effect at lower resolution, i.e. pixel-level. At chapter 6, we assessed the performance of stochastic search variable selection for the fixed effects of geostatistical models and compared three different Bayesian variable selection methods for spatially varying coefficients with a conditional autoregressive Gaussian process imposed on the varying part of the models. Additionally, we assessed the co-implementation of stochastic search variable selection with the selection processes for the spatially varying coefficients. Our work followed a recent renewed interest in Bayesian variable selection methods for mixed-effect spatial models. Our results are of great importance to the field of Bayesian modelling, as spatial models are computationally expensive and to this regard, selection methods that can reduce the parameter space and therefore the computation time for model fitting and prediction are of great interest. At Chapter 5, we provide a novel methodology to model the effect of a set of covariates on the changes in mortality risk between two survey time-points. A methodological issue was the difference in locations at which we had available data for the two time-points. We fitted and predicted the mortality risk at locations of the second survey and subsequently modelled the logarithmic difference of the Weibull scale parameters corresponding to each survey dataset. The presented methodology is valuable to modellers aiming to assess the impact of diseases or health interventions on the changes in mortality risk in Africa, as cross-sectional, survey data

constitute the primary source of information in the absence of improved vital registration systems.

7.2 Study limitations

The work presented in this PhD thesis was based on a number of assumptions and limitations that have, to some extent, influenced the results, and therefore the conclusions presented. One key point is that our work is based on statistical models that establish associations and do not imply causal relations. For instance, it is biologically proven that the malaria-anemia comorbidity leads to mortality and yet our analysis presented in Chapter 2 could not claim any causal associations. To establish such relations, a causal inference approach should have been rather used. We treated under-five mortality as the main outcome in most of our modelling efforts and it should be carefully noted that these survival data reflect a five years period preceding the time at which the DHS surveys took place. Contrarily, the biomarker tests from which we extracted the data concerning the childhood diseases, e.g. malaria and anemia, were performed only for the alive children exactly when the DHS surveys were conducted. The same issue concerns the health interventions considered in our studies. This assumption is particularly strong for data stemming from the most recent DHS surveys, when the scaling up of health interventions had already taken place. Further, the structure of the data rendered impossible to use individual level diseases or interventions data. Another limitation concerning Chapter 2 is that children with severe malaria anemia are less likely to participate in DHS or MIS, leading to underestimation of corresponding effect in our analysis. A crucial point is also that throughout our work we considered that climatic and environmental factors can be used as proxies of vector-borne diseases and that socio-economic status can account for poverty-related conditions, for instance malnutrition. The latter is a noteworthy assumption, as the WHO has reported that malnutrition is linked to 1 out of 3 under-five deaths. At Chapter 5, we assumed that environmental and climatic factors could solely explain all-cause, under-five mortality. This hypothesis was because we did not have available DHS data at locations of the second

surveys for the first time point and hence we relied solely on climatic data, which were available for all sites and time points. This was a rather strong assumption that might have led to underestimation of the effect of health interventions on changes in mortality risk. Concerning chapter 6, our analysis did not take into account methods for spatio-temporal varying coefficients. We also focused solely on CAR-structured varying coefficients while other specifications were also available and perhaps less parameter-heavy.

7.3 Future research recommendations

The malaria-anemia comorbidity indicator is a novel approach to measure malaria-related mortality but was only tested using all-cause mortality data from cross-sectional surveys. Future research should further evaluate the use of the comorbidity indicator in estimating deaths for under-five children using malaria-specific mortality data or information stemming from different sources than the DHS. Additionally, future studies should try to quantify the anemia attributed to malaria infection, as our study did not explore the causal agent for anemia in individuals classified to experience malaria-anemia comorbidity. Furthermore, further insights from the under-five mortality could be achieved by modelling childhood diseases with all-cause, under-five mortality. Causal modelling procedures for the effects of the leading causes of under-five mortality with fever would greatly improve our understanding of febrile response in Africa. As the introduction of a new health intervention could result in better equity outcomes, it is worth studying the potential contribution of the new malaria vaccine in health equity for Africans.

Bibliography

Alegana, V.A. *et al.* National and sub-national variation in patterns of febrile case management in sub-Saharan Africa. *Nat. Commun.* **9**, 4994 (2018).

Amir, A., Cheong, F.W., Ryan de Silva, J., Liew, J.W.K & Lau, Y.L. *Plasmodium knowlesi* malaria: current research perspectives. *Infect Drug Resist.* **11**, 1145–1155 (2018).

Arunda, M., Emmelin, A. & Asamoah, B.O. Effectiveness of antenatal care services in reducing neonatal mortality in Kenya: analysis of national survey data. *Glob. Health Action*. **10,** 1328796 (2017).

Ashley, E.A., Phyo, A.P. & Woodrow, C.J. Malaria. *Lancet* **391**, 1608–1621 (2018).

Azage, M., Kumie, A., Worku, A., Bagtzoglou, A.C. & Anagnostou, E. Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. *PLoS One* **12**, 0186933 (2017).

Banerjee, S., Carlin, B.P. & Gelfand, A.E. *Hierarchical Modeling and Analysis for Spatial Data, Second Edition*. (Chapman and Hall/CRC, Boca Raton, Florida, 2014).

Bangirana, P. *et al*. Severe malarial anemia is associated with long-term neurocognitive impairment. *Clin Infect Dis.* **59,** 336—344 (2014).

Barros, A.J.D. *et al*. Equity in maternal, newborn, and child health interventions in Countdown to 2015: a retrospective review of survey data from 54 countries. *Lancet*. **379,** 1225–1233 (2012).

Bartoloni, A. & Zammarchi, L. Clinical Aspects of Uncomplicated and Severe Malaria. *Mediterr J Hematol Infect Dis*. **4**, 2012026 (2012).

Bartram, J., Lewis, K., Lenton, R. & Wright, A. Focusing on improved water and sanitation

for health. *Lancet* **365**, 810–812 (2005).

Beck-Johnson, L.M. *et al.* The Effect of Temperature on Anopheles Mosquito Population Dynamics and the Potential for Malaria Transmission. *PLoS One* **8**, 79276 (2013).

Bhatt, S. *et al.* The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature* **526**, 207–211 (2015).

Bisoffi, Z. & Buonfrate, D. When fever is not malaria. *Lancet Glob Health* **1**, 11–12 (2013).

Boehm Vock, L., Reich, B., Fuentes, M. & Dominici, F. Spatial Variable Selection Methods for Investigating Acute Health Effects of Fine Particulate Matter Components. *Biometrics* **71**, 167–177 (2015).

Bouyou-Akotet, M.K. *et al*. Impact of Plasmodium falciparum infection on the frequency of moderate to severe anaemia in children below 10 years of age in Gabon. *Malar J.* **8,** 166 (2009).

Brown, P., Vannucci, M. & Fearn, T. Multivariate Bayesian variable selection and prediction. *J. Royal Stat. Soc. B* **60**, 627–641 (1998).

Brown, P. Model-Based Geostatistics the Easy Way. *J. Stat. Softw.* **63** (2015).

Bustreo, F., Okwo-Bele, J-M. & Kamara, L. World Health Organization perspectives on the contribution of the Global Alliance for Vaccines and Immunization on reducing child mortality. Arch. Dis. Child. **100**, 34–37 (2015).

Cai, B., Lawson, A.B., Hossain, M.M., Kirby, R.S. & Liu, J. Bayesian semiparametric model with spatially-temporally varying coefficients selection. *Stat. Med.* **32**, 3670–3685 (2013).

Carpenter, B. *et al.* Stan: A probabilistic programming language. *J. Stat. Softw*. **76**, 1–32 (2017).

CDC. *About Malaria. Biology*. https://www.cdc.gov/malaria/about/biology/index.html (2018a).

CDC. *Rapid Diagnostic Tests: How They Work*. https://www.cdc.gov/malaria/malaria_ worldwide/reduction/dx_rdt.html (2018b).

Ceesay, S.J. *et al*. Malaria Prevalence among Young Infants in Different Transmission Settings, Africa. *Emerg Infect Dis.* **21,** 1114—1121 (2015).

Chammartin, F. *et al.* Statistical methodological issues in mapping schistosomiasis survey data. *Acta Trop.* **128**, 345–352 (2013).

Chandramohan, D. *et al.* Should Data from Demographic Surveillance Systems Be Made More Widely Available to Researchers? *PLoS Med* **5**, 57 (2008).

Chen, Z. Random Effects Selection in Linear Mixed Models. *Biometrics* **59**, 762–769 (2003).

Choi, J. & Lawson, A.B. Bayesian spatially dependent variable selection for small area health modeling. *Stat. Methods Med. Res.* **27**, 234–249 (2016).

Choi, J. & Lawson, A.B. A Bayesian two-stage spatially dependent variable selection model for space–time health data. *Stat. Methods Med. Res.* **28**, 2570–2582 (2018).

Cordoba, M.A., Fuentes, M., Guinness, J. & Xie, L. Multivariate Spatial-Temporal Variable Selection with Applications to Seasonal Tropical Cyclone Modeling. https://arxiv.org/abs/1805.03318 (2018).

Cowman, A.F., Healer, J., Marapana, D. & Marsh, K. Malaria: biology and disease. *Cell* **167**, 610–624 (2016).

Cressie, N., Calder C.A., Clark, J.S., Hoef, J.M. & Wikle, C.K. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modelling. *Ecol App.* **19**, 553–570 (2009).

Crump, J.A. *et al.* Etiology of Severe Non-malaria Febrile Illness in Northern Tanzania: A Prospective Cohort Study. *PLoS Negl Trop Dis* **7**, 2324 (2013).

D'Acremont, V, Lengeler, C. & Genton, B. Reduction in the proportion of fevers associated with *Plasmodium falciparum* parasitaemia in Africa: a systematic review. *Malar. J* **9**, 240 (2010).

D'Acremont, V. *et al*. Beyond Malaria — Causes of Fever in Outpatient Tanzanian Children. *N Engl J Med* **370**, 809–817 (2014).

Dalrymple, U. *et al.* The contribution of non-malarial febrile illness co-infections to *Plasmodium falciparum* case counts in health facilities in sub-Saharan Africa. *Malar. J* **18**, 195 (2019).

Degarege, A., Fennie, K., Degarege, D., Chennupati, S. & Madhivanan, P. Improving

socioeconomic status may reduce the burden of malaria in sub Saharan Africa: A systematic review and meta-analysis. *PLoS One* **14**, 0211205 (2019).

Dellaportas, P., Forster, J. & Ntzoufras, I. On Bayesian model and variable selection using MCMC. ***Stat Comput* 12,** 27–36 (2002).

Dhaliwal, L.K. *Health Equity and Sustainable Development Goals: Role and the Complexity*. In: Leal Filho, W. *et al. Good Health and Well-Being. Encyclopedia of the UN Sustainable Development Goals.* (Springer, Cham, 2019).

Diggle, P.J., Tawn, J.A. & Moyeed, R.A. Model-based geostatistics. *Appl. Statist.* **47**, 299–350 (1998).

DOCTORS WITHOUT BORDERS. *DRC: Urgent Action Needed to Prevent Malaria Deaths in Orientale Province.* https://www.doctorswithoutborders.org/news-stories/press-release/drc-urgent-action-needed-prevent-malaria-deaths-orientale-province (2018).

Evans, S.S., Repasky, E.A. & Fisher, D.T. Fever and the thermal regulation of immunity: the immune system feels the heat. *Nat Rev Immunol.* **15**, 335–349 (2015).

Finley, A. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. *Methods Ecol. Evol.* **2**, 143–154 (2011).

Foote, E.M. *et al*. Determinants of Anemia among Preschool Children in Rural, Western Kenya. *Am J Trop Med Hyg.* **88,** 757—764 (2013).

Frühwirth-Schnatter, S. & Wagner, H. Bayesian variable selection for random intercept modeling of Gaussian and non-Gaussian data. *Bayesian Statistics* **9**, 165–200 (2010).

García-Díaz, P., Ross, J., Ayres, C. & Cassey, P. Understanding the biological invasion risk posed by the global wildlife trade: propagule pressure drives the introduction and establishment of Nearctic turtles. *Glob. Change Biol.* **21**, 1078–1091 (2015).

GBD 2016 Mortality Collaborators. Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. **390,** 1084–1150

(2017).

GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1859–1922 (2018).

Gelfand, A.E. & Smith, A.F.M. Sampling-Based Approaches to Calculating Marginal Densities. *J. Am. Stat. Assoc.* **85**, 398–409 (1990).

Gelfand, A., Kim, H.J., Sirmans, C. & Banerjee, S. Spatial modelling with spatially varying coefficient processes. *J. Am. Stat. Assoc.* **98**, 387–396 (2003).

Gelfand, A. & Vounatsou, P. Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* **4**, 11–25 (2003b).

Gelman, A. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* **1**, 515–534 (2006).

Gelman, A. *et al. Bayesian Data Analysis. Third edition.* (Chapman and Hall/CRC, Boca Raton, Florida, 2013).

Gemperli, A., Vounatsou, P., Sogoba, N. & Smith, T. Malaria mapping using transmission models: application to survey data from Mali. *Am J Epidemiol.* **163,** 289—297 (2006).

George, E. & McCulloch, R. Variable Selection via Gibbs sampling. *J. Am. Stat. Assoc* **88**, 881–889 (1993).

George, E. & McCulloch, R. Approaches for bayesian variable selection. *Statistica Sinica* **7**, 339–374 (1997).

Gething, P.W. *et al*. Mapping Plasmodium falciparum Mortality in Africa between 1990 and 2015. *N Engl J Med.* **375,** 2435–2445 (2016).

Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF. *Ghana Malaria Indicator Survey 2016*. (GSS, GHS, and ICF, Accra, Ghana, and Rockville, Maryland, USA, 2017).

Giardina, F. *et al.* Effects of vector-control interventions on changes in risk of malaria parasitaemia in sub-Saharan Africa: a spatial and temporal analysis. *Lancet Glob*

Health **2**, 601–615 (2014).

Golding, N. *et al.* Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *Lancet*. **390,** 2171–2182 (2017).

Green, P. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711–732 (1995).

Greenwell, F. & Neuman, M. *Children's anemia levels in West Africa: a good proxy for malaria morbidity?* http://paa2006.princeton.edu/papers/61268 (2018).

Haldar, K. & Mohandas, N. Malaria, erythrocytic infection, and anemia. *Hematology Am Soc Hematol Educ Program.* 87—93 (2009).

Hancock, P.A. *et al.* Associated patterns of insecticide resistance in field populations of malaria vectors across Africa. *PNAS* **115**, 5938–5943 (2018).

Hemingway, J. *et al*. Averting a malaria disaster: will insecticide resistance derail malaria control? *Lancet* **387**, 1785–1788 (2016).

Hoeting, J.A. The Importance of Accounting for Spatial and Temporal Correlation in Analyses of Ecological Data. *Ecol App.* **19**, 574–577 (2009).

ICF. *The Demographic and Health Surveys (DHS) Program*. https://www.icf.com/clients/dhs-program (2019).

Imdad, A. *et al.* Impact of vitamin A supplementation on infant and childhood mortality. *BMC Public Health*. **11,** S20 (2011).

INDEPTH Network. *Vision, Mission & Strategic Objectives*. http://www.indepth-network.org/about-us/vision-mission-strategic-objectives-0 (2019).

Kai, O.K. & Roberts, D.J. The pathophysiology of malarial anaemia: where have all the red cells gone? *BMC Med.* **6,** 24 (2008).

Kateera, F. *et al*. Malaria, anaemia and under-nutrition: three frequently co-existing conditions among preschool children in rural Rwanda. *Malar J.* **14,** 440 (2015).

Kazembe, L.N., Appleton, C.C. & Kleinschmidt I. Spatial analysis of the relationship between early childhood mortality and malaria endemicity in Malawi. *Geospat Health.* **2,** 41—50 (2007).

Khagayi, S., Amek, N., Bigogo, G., Odhiambo, F. & Vounatsou, P. Bayesian spatio-temporal modeling of mortality in relation to malaria incidence in Western Kenya. *PLoS ONE* **12,** e0180516 (2017).

Khagayi, S. *et al.* Modelling the relationship between malaria prevalence as a measure of transmission and mortality across age groups. *Malar J.* **18**, 247 (2019).

Kinney, S. & Dunson, D. Fixed and Random Effects Selection in Linear and Logistic Models. *Biometrics* **63**, 690–698 (2007).

Kleinschmidt, I. Implications of insecticide resistance for malaria vector control with long-lasting insecticidal nets: a WHO-coordinated, prospective, international, observational cohort study. *Lancet Infect Dis* **18**, 640–649 (2018).

Kuo, L. & Mallick, B. Variable selection for regression models. *Sankhya Ser. B* **60**, 65–81 (1998).

Leviton, A. Definitions of attributable risk. *Am. J. Epidemiol.* **98**, 231 (1973).

Leuenberger, A. *et al*. Health impact assessment and health equity in sub-Saharan Africa: A scoping review. *Environ. Impact Assess. Rev*. **79,** 106288 (2019).

Lindgren, F., Rue, H. & Lindstrom, J. An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach (with discussion). *J R Stat Soc Series B Stat Methodol.* **73**, 423–498 (2011).

Liu, L. *et al*. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet*. **388,** 3027–3035 (2016).

Liu, Y. *et al.* Impact of meteorological factors on lower respiratory tract infections in children. *J Int Med Res.* **44**, 30–41 (2016).

Lo, N.C. *et al.* State of deworming coverage and equity in low-income and middle-income countries using household health surveys: a spatiotemporal cross-sectional study. *Lancet Glob Health* **7**, 1511–1520 (2019).

Magalhães, R.J.S. *et al*. Role of malnutrition and parasite infections in the spatial variation in children's anaemia risk in northern Angola. *Geospat Health.* **7,** 341—354 (2013).

Maze. M.J. *et al.* The epidemiology of febrile illness in sub-Saharan Africa: implications for diagnosis and management. *Clin Microbiol Infect* **24**, 808–814 (2018).

McCuskee, S., Brickley, E.B., Wood, A. & Mossialos, E. Malaria and Macronutrient Deficiency as Correlates of Anemia in Young Children: A Systematic Review of Observational Studies. *Ann Glob Health* **80,** 458—465 (2014).

McElreath, R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan.* (Chapman and Hall/CRC, Boca Raton, Florida, 2014).

McGovern, M.E. & Canning, D. Vaccination and All-Cause Child Mortality From 1985 to 2011: Global Evidence From the Demographic and Health Surveys. *Am J Epidemiol.* **182,** 791–798 (2015).

Menon, M.P. & Yoon, S.S. Prevalence and Factors Associated with Anemia among Children under 5 Years of Age—Uganda, 2009. *Am J Trop Med Hyg.* **93,** 521—526 (2015).

Millogo, O., Doamba, J.E.O., Sié, A., Utzinger, J. & Vounatsou, P. Geographical variation in the association of child, maternal and household health interventions with under-five mortality in Burkina Faso. *PloS one*. **14,** e0218163 (2019).

Mirsaeidi, M. *et al.* Climate Change and Respiratory Infections. *Ann Am Thorac Soc.* **13**, 1223–1230 (2016).

Moraleda, C. *et al*. Anaemia in hospitalised preschool children from a rural area in Mozambique: a case control study in search for aetiological agents. *BMC Pediatr.* **17,** 63 (2017).

Morrison, J. *et al*. Exploring the equity impact of a maternal and newborn health intervention: A qualitative study of participatory women's groups in rural South Asia and Africa. *Int J Equity Health*. **18,** 55 (2019).

Mosser, J.F., *et al.* Mapping diphtheria-pertussis-tetanus vaccine coverage in Africa, 2000–2016: a spatial and temporal modelling study. *Lancet* **393**, 1843–1855 (2019).

Muoneke, V.U., Ibekwe, R.C., Nebe-Agumadu, H.U. & Ibe, B.C. Factors associated with mortality in under-five children with severe anemia in Ebonyi, Nigeria. *Indian Pediatr.* **49,** 119—123 (2012).

Murray, C.J. *et al*. Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet.* **379,** 413—431 (2012).

Nambuusi, B.B. Assessing the effects of interventions on child and maternal health-related outcomes in Uganda. (Doctoral Thesis, University of Basel, Basel, Switzerland, 2019).

Newcombe, P. *et al.* Weibull regression with Bayesian variable selection to identify prognostic tumour markers of breast cancer survival. *Stat. Methods Med. Res.* **26**, 414–436 (2014).

Obonyo, C.O., Vulule, J., Akhwale, W.S. & Grobbee, D.E. In-hospital morbidity and mortality due to severe malarial anemia in western Kenya. *Am J Trop Med Hyg.* **77,** 23—28 (2007).

O'Brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Qual Quant.* **41,** 673—690 (2007).

Ogoina, D. Fever, fever patterns and diseases called 'fever' – A review. *J. Infect. Public Health* **4**, 108–124 (2011).

O'Hara, R. & Sillanpää, M. A review of Bayesian Variable Selection Methods: What, How and Which. *Bayesian Anal.* **4**, 85–117 (2009).

O'Meara, W.P. *et al.* Etiology of Pediatric Fever in Western Kenya: A Case–Control Study of Falciparum Malaria, Respiratory Viruses, and Streptococcal Pharyngitis. *Am. J. Trop. Med. Hyg.* **92**, 1030–1037 (2015).

Osterbauer, B. *et al*. Factors associated with malaria parasitaemia, malnutrition, and anaemia among HIV-exposed and unexposed Ugandan infants: a cross-sectional survey. *Malar J.* **11,** 432 (2012).

Papaioannou, I., Utzinger, U., & Vounatsou, P. Malaria-anemia comorbidity prevalence as a measure of malaria-related deaths in sub-Saharan Africa. *Sci. Rep.* **9,** 11323 (2019).

Patz, J.A. & Olson S.H. Malaria risk and temperature: Influences from global climate change and local land use practices. *PNAS* **103**, 5635–5636 (2006).

Pedersen, U.B. *et al.* Comparison of the spatial patterns of schistosomiasis in Zimbabwe

at two points in time, spaced twenty-nine years apart: is climate variability of importance? *Geospat. Health* **12**, 505 (2017).

Peng, B. *et al.* An Integrative Framework for Bayesian Variable Selection with Informative Priors for Identifying Genes and Pathways. *PLoS One* **8**, 67672 (2013).

Perkins, D.J. *et al*. Severe malarial anemia: innate immunity and pathogenesis. *Int J Biol Sci.* **7,** 1427—1442 (2011).

Plummer, M. *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.* (Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria, 2003).

Polson, N.G. & Scott, J.G. On the Half-Cauchy Prior for a Global Scale Parameter. *Bayesian Anal.* **7**, 887–902 (2012).

Prasad, N., Murdoch, D.R., Reyburn, H. & Crump, J.A. Etiology of Severe Febrile Illness in Low- and Middle-Income Countries: A Systematic Review. *PLoS One* **10**, 0127962 (2015).

R Core Team (2019). *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing, Vienna, Austria, 2019).

Reich, B., Fuentes, M., Herring, A. & Evenson, K. Bayesian Variable Selection for Multivariate Spatially Varying Coefficient Regression. *Biometrics* **66**, 772–782 (2010).

Reunanen, J. Overfitting in making comparisons between variable selection methods. *J. Mach. Learn. Res.* **3**, 1371–1382 (2003).

Ricotta, E.E, Frese, S.A., Choobwe, C., Louis, T.A. & Shiff C.J. Evaluating local vegetation cover as a risk factor for malaria transmission: a new analytical approach using ImageJ. *Malar. J.* **13**, 94 (2014).

Righetti, A.A. *et al*. Dynamics of Anemia in Relation to Parasitic Infections, Micronutrient Status, and Increasing Age in South-Central Côte d'Ivoire. *J Infect Dis.* **207,** 1604—1615 (2013).

Rogers, K. *Parasitic disease*. https://www.britannica.com/science/parasitic-disease (Encyclopædia Britannica, inc., 2017).

RStudio Team (2018). *RStudio: Integrated Development for R.* (RStudio, Inc., Boston, MA, 2018).

Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J R Stat Soc Series B Stat Methodol.* **71,** 319—392 (2009).

Rumisha, S.F., Smith, T.A., Masanja, H., Abdulla, S. & Vounatsou, P. Relationship between child survival and malaria transmission: an analysis of the malaria transmission intensity and mortality burden across Africa (MTIMBA) project data in Rufiji demographic surveillance system, Tanzania. *Malar J*. **13,** 124 (2014).

Rutstein, S.O. & Johnson, K. *The DHS Wealth Index. DHS Comparative Reports No. 6.* (ORC Macro, Calverton, Maryland, 2004).

Sinka, M.E. *et al*. The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis. *Parasit Vectors* **3**, 117 (2010).

Snow, R.W. *et al.* Changing Malaria Prevalence on the Kenyan Coast since 1974: Climate, Drugs and Vector Control. *PLoS ONE* **10**, 0128792 (2015).

Solar, O. & Irwin, A. *A conceptual framework for action on the social determinants of health. Social Determinants of Health Discussion Paper 2 (Policy and Practice)*. (World Health Organization, Geneva, 2010).

Shoo, R. *Reducing Child Mortality: The Challenges in Africa.* (World Health Organization, Geneva, 2007).

Soulard, V. *et al*. *Plasmodium falciparum* full life cycle and *Plasmodium ovale* liver stages in humanized mice. *Nat. Commun*. **6**, 7690 (2015).

Srivastana, S. & Chen, L. Comparison between the stochastic search variable selection and the least absolute shrinkage and selection operator for genome-wide association studies of rheumatoid arthritis. *BMC Proc.* **3**, S21 (2009).

Ssempiira, J. *et al.* Interactions between climatic changes and intervention effects on malaria spatio-temporal dynamics in Uganda. *Parasite Epidemiol Control* **3**, 00070 (2018).

Stan Development Team. 2018. *RStan: the R interface to Stan.* R package version 2.17.3. http://mc-stan.org.

Stevens, G.A. *et al.* Trends and mortality effects of vitamin A deficiency in children in 138 low-income and middle-income countries between 1991 and 2013: a pooled analysis of population-based surveys. *Lancet Glob Health* **3**, 528–536 (2015).

Taylor, C., Florey, L. & Ye, Y. Equity trends in ownership of insecticide-treated nets in 19 sub-Saharan African countries. *Bull World Health Organ* **95**, 322–332 (2017).

Tekelab, T., Chojenta, C., Smith, R. & Loxton, D. The impact of antenatal care on neonatal mortality in sub-Saharan Africa: A systematic review and meta-analysis. *PLoS ONE* **14**, 0222566 (2019).

Teng, M., Nathoo, F.S. & Johnson, T.D. Bayesian analysis of functional magnetic resonance imaging data with spatially varying auto-regressive orders. *J. Royal Stat. Soc. C.* **68**, 521–541 (2018).

The DHS program. *DHS Overview*. https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm (2019a).

The DHS program. *Using Biormakers to Collect Health Data*. https://dhsprogram.com/What-We-Do/Biomarkers.cfm (2019b).

The Malaria Atlas Project. *Mosquito Malaria Vectors*. https://map.ox.ac.uk/mosquito-malaria-vectors/ (2019).

Tüchler, R. Bayesian variable selection for logistic models using auxiliary mixture sampling. *J. Comput. Graph. Stat.* **17**, 76–94 (2008).

Tusting, L.S. *et al.* The evidence for improving housing to reduce malaria: a systematic review and meta-analysis. *Malar. J.* **14**, 2019 (2015).

Tuteja, R. Malaria – an overview. *FEBS J.* **274**, 4670–4679 (2007).

UNICEF. *Committing to Child Survival: A Promise Renewed. Progress Report 2014*. (United Nations Children's Fund, New York, 2014).

UNICEF, WHO, World Bank, UN DESA/Population Division. *Levels and trends in Child Mortality 2018.* (United Nations Children's Fund, New York, 2018).

UNICEF. *Convention on the Rights of the Child.* https://www.unicef.org/child-rights-convention/convention-text (2019).

UN. Department of Economic and Social Affairs. *Sustainable Development Goals.* https://sustainabledevelopment.un.org/topics/sustainabledevelopmentgoals (2018).

UN. *Goal 4: Reduce child mortality.* https://www.un.org/millenniumgoals/childhealth.shtml (2019a).

UN. Sustainable Development Goals. Knowledge Platform. *Sustainable Development Goal 3.* https://sustainabledevelopment.un.org/sdg3 (2019b).

United Nations Inter-agency Group for Child Mortality Estimation (UN IGME). *Levels & Trends in Child Mortality: Report 2019, Estimates developed by the United Nations Inter-agency Group for Child Mortality Estimation*. (United Nations Children's Fund, New York, 2019).

United Nations. Department of Economic and Social Affairs. Population Division. *Population Facts No. 2017/6, December 2017 - Africa unlikely to achieve SDG target on child mortality without unprecedented effort to accelerate progress.* (United Nations, New York, 2017a).

United Nations. *The Millennium Development Goals Report 2015.* (United Nations, New York, 2015).

USAID. *The Demographic and Health Surveys Program*. https://www.usaid.gov/what-we-do/global-health/cross-cutting-areas/demographic-and-health-surveys-program (2018).

Wagner, H. & Duller, C. Bayesian model selection for logistic regression models with random intercept. *Comput. Stat. Data Anal.* **56**, 1256–1274 (2012).

Wariri, O. *et al*. Tracking coverage, dropout and multidimensional equity gaps in immunisation systems in West Africa, 2000–2017. *BMJ global health* **4**, 001713 (2019).

Watanabe, S. Asymptotic Equivalent of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *J. Mach. Learn. Res.* **11**, 3571–

3594 (2010).

Wiebe, A. *et al.* Geographical distributions of African malaria vector sibling species and evidence for insecticide resistance. *Malar. J.* **16**, 85 (2017).

Whitlock, E. *et al.* Postoperative Delirium in a Substudy of Cardiothoracic Surgical Patients in the BAG-RECALL Clinical Trial. *Anesth. Analg.* **118**, 809–817 (2014).

WHO. *How malaria RDTs work*. https://www.who.int/malaria/areas/diagnosis/rapid-diagnostic-tests/about-rdt/en/ (2015a).

WHO. *GLOBAL TECHNICAL STRATEGY FOR MALARIA 2016–2030.* (World Health Organization, Geneva, 2015b).

WHO. *World Malaria Report 2017.* (World Health Organization, Geneva, 2017a).

WHO. *WHO Malaria Policy Advisory Committee (MPAC) meeting. Meeting Report.* (World Health Organization, Geneva, 2017b).

WHO. *Overview of malaria treatment*. https://www.who.int/malaria/areas/treatment/overview/en/ (2018a).

WHO. *World Malaria Report 2018*. (World Health Organization, Geneva, 2018b).

WHO. *The top 10 causes of death*. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (2018c).

WHO. *WHO Malaria Policy Advisory Committee (MPAC) meeting. Meeting Report.* (World Health Organization, Geneva, 2018d).

WHO. *Rapid diagnostic tests*. https://www.who.int/malaria/areas/diagnosis/rapid_diagnostic_tests/en/ (2019a).

WHO. *Q&A on the malaria vaccine implementation programme (MVIP).* https://www.who.int/malaria/media/malaria-vaccine-implementation-qa/en/ (2019b).

WHO. *Under-five mortality*. https://www.who.int/gho/child_health/mortality/mortality_under_five_text/en/ (2019c).

WHO. *Vitamin A supplementation in infants and children 6–59 months of age.* https://www.who.int/elena/titles/vitamina_children/en/ (2019d).

WHO. *Environmental health in emergencies. Vulnerable groups.* https://www.who.int/

environmental_health_emergencies/vulnerable_groups/en/ (2019e).

WHO. Regional office for Africa. *Child health*. https://www.afro.who.int/health-topics/child-Health (2019f).

Woods, C.R., Davis, D.W., Duncan, S.D. Myers, J.A. & O'Shea, T.M. Variation in classification of live birth with newborn period death versus fetal death at the local level may impact reported infant mortality rate. *BMC Pediatr.* **14**, 108 (2014).

Xu, S. Estimating polygenic effects using markers of the entire genome. *Genetics* **163**, 789–801 (2003).

Yang, M. Bayesian variable selection for logistic mixed model with nonparametric random effects. *Comput. Stat. Data Anal.* **56**, 2663–2674 (2011).

Yaya, S., Bishwajit, G., Okonofua, F. & Uthman, O.A. Under five mortality patterns and associated maternal risk factors in sub-Saharan Africa: A multi-country analysis. *PLoS One* **13**, 0205977 (2018).

Ye, Y. Wamukoya, M., Ezeh, A., Emina, J.B.O. & Sankoh, O. Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Sahara Africa? *BMC public health* **12**, 741 (2012).

Yourkavitch, J., Burgert-Brucker, C., Assaf, S. & Delgado, S. Using geographical analysis to identify child health inequality in sub-Saharan Africa. *PLoS One* **13**, 0201870 (2018).

Zhang, F., Jiang, W., Wong, P. & Wang, J.P. A Bayesian probit model with spatially varying coefficients for brain decoding using fMRI data. *Stat. Med.* **35**, 4380–4397 (2016).

Zilversmit, M. & Perkins, S. *Plasmodium*. http://tolweb.org/Plasmodium/68071 (2019).

Curriculum Vitae

Professional experience

02/2020 – now	**Consultant – PriceWaterhouseCoopers.** Data science consultant for the Pharma/Life Science industries.

Education

09/2015 – 01/2020	**PhD in Epidemiology – Biostatistics Unit. Swiss Tropical and Public Health Institute,** Basel, Switzerland. Project: *"Innovative modelling to optimize control of childhood anaemia across Africa"* (ERC). Advisors: Dr Penelope Vounatsou and Prof. Jürg Utzinger.
09/2014 – 08/2015	**MSc in Statistics – School of Mathematics. The University of Manchester,** Manchester, UK. Dissertation: *"Analysis of a cancer drug combination study"*. Advisor: Dr Alexander Donev. Grade: Distinction.
09/2009 – 08/2014	**BSc in Mathematics – Department of Mathematics. Aristotle University of Thessaloniki**, Thessaloniki, Greece. Grade: Merit.
2009	High School Diploma. 14th Public school of Thessaloniki, Greece. Grade: Distinction.

Research Interests

Bayesian modelling, variable selection, malaria, spatial statistics and Bayesian computation.

Publications

Papaioannou, I., Utzinger, J. & Vounatsou, P. Malaria-anemia comorbidity prevalence as a measure of malaria-related deaths in sub-Saharan Africa. *Sci. Rep.***9**, 11323 (2019).

Computer skills

Advanced user of R programming language and Python, experienced in doing Bayesian data analysis with STAN/INLA/JAGS/BUGS/NIMBLE